Evaluation of an automatic lean meat percentage quantification method based on a partial volume model from computed tomography scans

Pau Xiberta[a,*], Anton Bardera[a], Imma Boada[a], Marina Gispert[b], Albert Brun[b], Maria Font-i-Furnols[b]

[a] *Graphics and Imaging Laboratory, Universitat de Girona, 17003 Girona, Catalonia.*
[b] *IRTA-Food Industries, Finca Camps i Armet, E-17121 Monells, Girona, Catalonia.*

Abstract

The quality of a pig carcass is mainly measured by the lean meat percentage (LMP), which can be virtually estimated from computed tomography (CT) scans. Different strategies exist to classify the CT voxels into tissues such as fat, lean and bone, being the thresholding-based methods the most commonly used. However, these methods are usually affected by the partial volume effect, and also by data variability, which is implicit from different CT scanners and protocols, since no standard behaviour has been defined. The aim of this paper is to extend an LMP quantification method which uses a partial volume model by adding a new step to detect the animal skin, and thoroughly evaluate the new approach by analysing each of its steps. The evaluation is performed by comparing the whole pipeline of the proposed approach with a simple thresholding method and a thresholding method with bone filling and skin detection, which is an intermediate step of the new

*Corresponding author
Email addresses: pau.xiberta@udg.edu (Pau Xiberta), anton.bardera@udg.edu (Anton Bardera), imma.boada@udg.edu (Imma Boada)

pipeline. Five experiments have been designed to test how accurate are the results of the method regarding the LMP values computed from the manual dissection, as well as the robustness to data variability. Two different manual dissection methodologies have been tested: the partial dissection, which estimates the LMP using the lean of the four main cuts of the carcass plus the tenderloin, and the total dissection, which uses the lean of the twelve main cuts. A total of 146 half carcasses have been used for this study (105 using the partial dissection methodology, and 41 using the total dissection one). To evaluate the experiments, the LMP values virtually obtained from the three methods have been compared mostly with the LMP values from the manual dissection, computing the coefficient of determination R^2 from the correlations, as well as the root mean square error of prediction by means of leave-one-out cross-validation. A statistical analysis is performed to resolve if two correlations are significantly different. The experiments' results confirm the high accuracy of the proposed approach for the LMP estimation, and mainly its high robustness to data variability. The experiments also disclose that the detection of the animal skin and its classification as a new tissue, instead of classifying it as lean, improve the results. The evaluated method has demonstrated to be as effective as the thresholding method with bone filling and skin detection, and more robust to data variability than the other evaluated methods.

Keywords: Lean meat percentage, Computed tomography, Partial volume effect, Segmentation, Pig carcass quantification

---

1. Introduction

Lean meat percentage (LMP) is a key parameter to measure pig carcass quality, it is compulsory in the Europe Union and it determines the basis for the price of the carcass. To compute the LMP from computed tomography (CT) scans, special methods to classify CT voxels into tissues according to its Hounsfield Unit (HU) values are required. Unfortunately, variability between animals and breeds, and also between scanners and protocols makes the definition of a standard correspondence between HU values and tissues difficult (Olsen et al. (2017)), and each country has defined its own model (Romvári et al. (2006); Font-i-Furnols et al. (2009); Daumas and Monziols (2011)). Moreover, the partial volume effect further complicates LMP computation, that is, voxels which are usually placed in the border between two tissue regions may have a big uncertainty, and they cannot be classified because they contain more than one tissue. This difficulty has been studied in other fields such as oncology (see Cysouw et al. (2017) for a review), but mainly in the field of neuroimaging (see Tohka (2014) for a review), evaluating its impact (Dukart and Bertolino (2014)), compiling different methods to enhance the image visualisation (Salminen et al. (2016)), and still proposing novel techniques to reduce the effect (Bural et al. (2015); Şener et al. (2016)).

To tackle the partial volume problem, different strategies have been proposed. Assuming a uniform probability for the non-pure tissues over the image, i.e. each partial volume voxel has the same probability for every non-pure tissue, Santago and Gage (1993) propose a model with six Gaussian distributions, three for the pure tissues and three for the two-class partial volume ones, with a set of parameters which have to be minimised to fit the

model to the histogram. With the same assumption, Laidlaw et al. (1998) reconstruct a continuous function incorporating neighbouring voxels information into the classification process to improve its accuracy, and Ruan et al. (2000) first use a mixture model to define a Gaussian distribution for each pure and partial volume tissue, and then reclassify the partial volume classes into the pure ones using a Markov random field and multifractal analysis.

Other studies assume little variation in the probability for the non-pure tissues between neighbouring voxels, which can be modelled using a Markov random field. Choi et al. (1991) use a maximum *a posteriori* estimation of partial volume voxels in multichannel images, and a method to iteratively reestimate the mean intensities of each tissue class in each slice, while Pham and Prince (2000) propose a similar method for single-channel images using a Bayesian approach which places a prior probability model on the parameters. Finally, Nocera and Gee (1997) describe a segmentation algorithm which also uses a maximum *a posteriori* estimation with an adaptive Bayesian approach, and takes into account both the partial volume and the shading effect.

Focusing on the LMP computation, several methods have been presented in the literature. Gangsei et al. (2016) and Jansons et al. (2016) use optical probes to collect certain variables, which is an efficient method when working with carcasses, and in Dobrowolski et al. (2004), Judas et al. (2007) and Font-i-Furnols et al. (2009) data from CT images is analysed using partial least squared regression, which does not require the classification of voxels in lean or fat. In this case, volume associated to each HU value is obtained from CT images and used as predictors in the regression. To build their regression equations, Kremer et al. (2013) and Bernau et al. (2015) use linear

traits measured by dual energy X-ray absorptiometry (DXA), while Lisiak et al. (2015) proposes a simpler approach using linear measurements over the carcass which do not need the use of expensive classification equipment. Another common method is to use thresholding techniques based on the HU values (Daumas and Monziols (2011)), and even mixing thresholding techniques with some manual interaction in a semi-automatic method (Bernau et al. (2015)). Kongsro et al. (2008) have applied a tresholding approach using lamb meat as well, and Lee et al. (2015) have adopted a similar method using beef, the latter also using a chemical analysis to compare the results with the thresholding method. To avoid dealing with the partial volume effect when using the thresholding techniques, some strategies have been proposed. In Vester-Christensen et al. (2009) the partial volume effect has been minimised applying a Bayesian 2D contextual classification scheme to classify voxels into fat, lean and bone. Differently, in Bardera et al. (2014) a five-step process which automatically quantifies fat, lean, and bone tissues from CT scans using a partial volume model based on the one presented by Van Leemput et al. (2003) is described, and a first validation of the method considering 10 carcasses is carried out.

The aim of this paper is to evaluate the quantification method presented in Bardera et al. (2014) considering 146 half carcasses which have been manually dissected after scanning (105 using a partial dissection, and 41 using a total dissection). The introduction to the method's pipeline of a new step which identifies and classifies the animal skin tissue is also analysed. The obtained results are compared in terms of LMP accuracy and robustness to data variability, and the importance and need for each step of the new

pipeline is discussed.

2. Materials and methods

2.1. Carcasses and computed tomography scanning

A total of 146 left half carcasses have been used for this study. From these, 133 carcasses come from two commercial abattoirs and have been selected to mimic the Spanish pig carcass population in terms of fat thickness, being all the three sexual types represented. These carcasses also come from several producers and commercial genotypes. Additionally, 13 carcasses from gilts, slaughtered at the pilot abattoir placed at IRTA-Monells, have also been used in this study. These carcasses are from 3 different genotypes as described in Carabús et al. (2014) and Font-i-Furnols et al. (2015). In total, carcasses included in this study have a carcass weight of 86.7 $\pm$ 8.7 kg, a fat thickness of 15.7 $\pm$ 3.8 mm measured at 6 cm of the midline between the 3rd and the 4th last ribs, and they are from three sexual types (47% females, 41% entire males and 12% castrated males). The Commission Delegated Regulation (EU) 2017/1182 (The European Commission (2017)) established a minimum of 120 carcasses representative of the population to be involved in a dissection trial. For this reason, the number and type of carcasses considered in this work is suitable to be used to evaluate the methodology proposed in this paper to determine carcass lean meat content.

At 24-48 h post mortem carcasses were CT scanned with a General Electric HiSpeed Zx/I device placed at IRTA-Monells. Acquisition parameters were those established by Font-i-Furnols et al. (2009) in carcasses evaluation, that is, 140 kV, 145 mA, Display Field of View (DFOV) between 460 and 500

mm, and matrix size 512×512 pixels. Images were acquired helically every 10 mm with pitch 1. Thus, there was not overlapping between images and all the carcasses were scanned completely.

2.2. Manual dissection

After scanning, carcasses were cut following the Walstra and Merkus method (Walstra and Merkus (1996)) and dissected by trained butchers. A total of 105 carcasses were dissected using the partial dissection methodology, i.e. the lean from the four main cuts (ham, shoulder, belly and loin) was manually separated with a knife and weighed. The LMP values were obtained dividing the weight of the lean of the four main cuts plus the tenderloin by the total weight of the four main cuts plus the tenderloin. A correction factor of 0.89 was applied to obtain the LMP values of the carcasses from these cuts, according to the European Regulation definition (The Commission of the European Communities (2008)). The other 41 carcasses were totally dissected, i.e. the lean of all the 12 cuts was manually obtained and weighed, and this weight was divided by the weight of the carcass to obtain the LMP (The Commission of the European Communities (2008)).

2.3. Automatic LMP quantification method based on a partial volume model

The proposed approach to quantify fat, lean and bone from CT carcasses is an improvement of the method presented in Bardera et al. (2014). We propose to extend this automatic five-step method with a new step which detects the animal skin. The six steps are illustrated in Figure 1 and described below. For more details see Bardera et al. (2014).

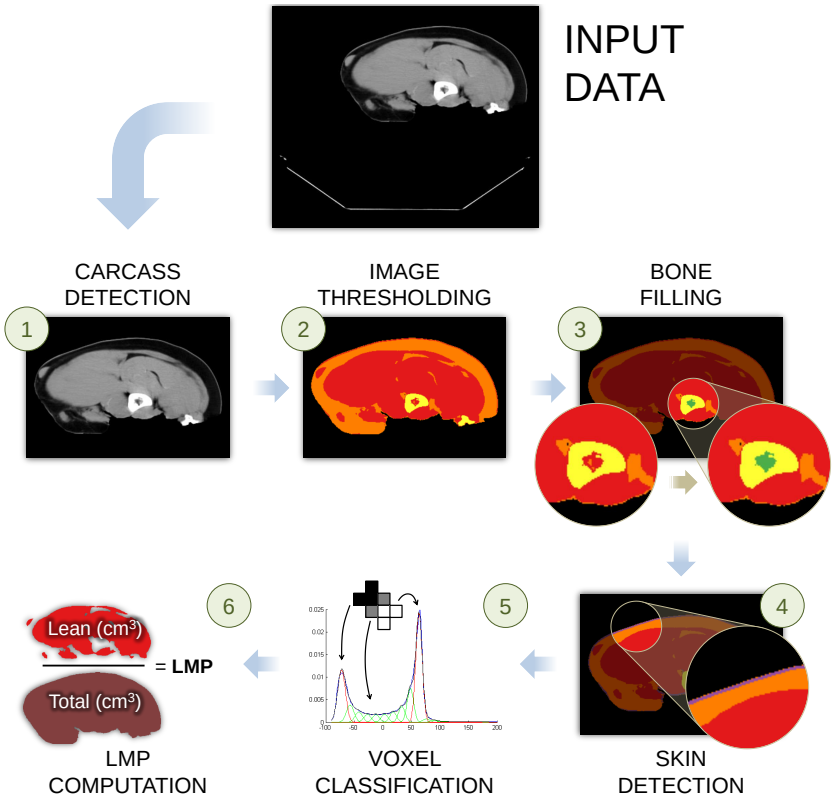


Figure 1: The six main steps of the proposed approach to compute the LMP values from pig carcasses CT images.

1. **Carcass detection.** The pig carcass is detected from the input CT scans, and other structures of the image such as the scanning table and the air are removed. Taking into account that the carcass lies over a cushion which has intensity values similar to those of the air, cropping the bottom part of the image is enough to remove the table and other supporting elements. Then, the carcass is only surrounded by voxels with very low intensities (air and cushion), and it can be detected using a simple thresholding method.

2. **Image thresholding.** A thresholding technique based on the HU values is applied in order to pre-identify fat, lean and bone tissues. By default, a value of -100 HU has been defined as a threshold between air/background and fat, 0 HU between fat and lean, and 120 HU between lean and bone (Font-i-Furnols et al. (2009)).
3. **Bone filling.** Since the marrow tissue is often confused with fat in the thresholding step, the marrow surrounded by bone tissue is also considered as bone. To achieve this result, a binary hole filling operation is carried out from the 2D bone mask obtained in the previous step.
4. **Skin detection.** Although it represents a little part of the whole carcass, the skin tissue should not be confused with the lean tissue, as both tissues have similar HU values. Knowing that the skin is the outermost tissue, and that the subcutaneous fat separates it from the lean tissue, a measure to detect the skin voxels is proposed, avoiding by this way to take them into account when computing the LMP values. All the voxels with values lying in the HU range of lean tissue are filtered so that the ones at 3 mm from the background are considered as skin. The distance is computed using a background binary mask obtained from the carcass detection in the first step, and a distance filter which measures the Euclidean distance of each voxel of the mask to the background. Once the distance is computed, a filtering process is applied in order to keep only those lean voxels obtained in the second step whose distance to the background is 3 mm or less, taking into account the size of a voxel, which in this case approaches 1 mm^3. This is the new step introduced in the method with respect to the one proposed

by Bardera et al. (2014).

5. **Pure class identification and partial volume model.** The pure and partial volume voxels are classified using the partial volume model without spatial correlation proposed by Van Leemput et al. (2003), which includes an iterative expectation-maximisation algorithm. In this case, the input of the method is just the histogram, leading to a very fast process, and the output is the probability of each intensity value to belong to the fat, lean or bone tissues. To compute the amount of voxels for the lean tissue, for example, all the voxels' probabilities of belonging to the lean tissue are added up, resulting in the estimated total number of lean voxels. Thus, a single voxel can contribute to the sum of different tissues. Figure 2 shows an original CT image and the partial volume classification of its voxels, including a separate image for each tissue indicating the probability of each voxel of belonging to it. Note the partial volume effect between the background and the already removed skin tissue in the carcass border of the fat voxels classification image.

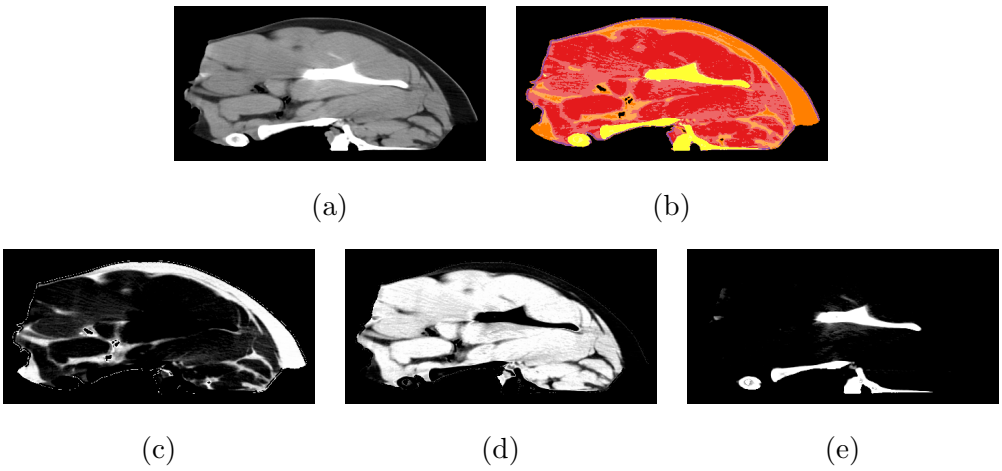


Figure 2: Partial volume classification of a particular CT image: (a) original CT image; (b) pure tissues (probability $> 90\%$, strong colours) and partial volume voxels ($50\% <$ probability $< 90\%$, pale colours), where orange = fat, red = lean, yellow = bone, and purple = skin (skin is not considered in the partial volume classification); (c, d, e) probability of each voxel of belonging to the fat, lean and bone tissues, respectively (white = 100% probability, black = 0% probability).

6. **LMP computation.** The volume of the lean meat obtained from the partial volume classification is divided by the volume of the whole carcass, including all the tissues. The result is the lean meat percentage, that is, the value of interest. To compute the volume, the corresponding number of voxels is multiplied by the volume of a single voxel.

2.4. Software implementation

The proposed approach has been implemented using C++, Qt, the Insight Toolkit (ITK) and the Visualisation Toolkit (VTK) libraries as a new module of the VisualPork software (Bardera et al. (2012)). VisualPork is an in-house software which supports DICOM standard and IHE profiles, integrates image

processing techniques, and provides 2D and 3D visualisation functionalities. It has been developed at the Graphics and Imaging Laboratory (GILAB) from the University of Girona in collaboration with experts from IRTA-Food Industries.

2.5. Evaluation metrics

The goal of the proposed approach is to find a method which is able to compute the LMP values from a pig carcass, with the purpose of getting a value as close as possible to the manually computed LMP. Thus, the first measure to be taken into account when analysing the CT images from the carcasses is the LMP, which is the ratio of the lean meat voxels to the total.

To compare the virtually obtained LMP values with the manually obtained ones, a correlation between these values is needed. Indeed, this correlation is needed to compare between two different methods, and also to compare between different ways of using the same method. The coefficient of determination R^2 (i.e. the square of the correlation coefficient R) will be used to analyse the correlations.

Considering the LMP values obtained from the manual dissection as the true value, the root mean square error of prediction (RMSEP) can also be computed by means of leave-one-out cross-validation as another measure of accuracy.

Finally, to determine whether the differences between the correlations of two different methods are significant, several statistical tests for comparing two correlations based on dependent groups with overlapping variables can be applied, being one of the most representative the one proposed by Steiger (1980). To apply these tests, the tool implemented by Diedenhofen and

Musch (2015) will be used with a significance level of 0.05.

2.6. Experiments' description

To evaluate the proposed approach and other related methods, different experiments have been carried out to determine whether the proposed approach can be selected as a method of reference to compute the LMP values of pig carcasses. Each experiment proposes alternatives of the pipeline described in Figure 1, and they are compared at least to the manual dissection results. Figure 3 shows a diagram where the alternatives of the pipeline are represented, always taking into account the 6 steps from the original pipeline. The experiments have been designed to evaluate the importance of each one of these steps. As previously mentioned, two different methodologies have been used to perform the manual dissection: the partial dissection, and the total dissection. All the experiments have been executed separately, using a different set of carcasses for each situation.

Experiment 1. In the first experiment, three methods to compute the LMP values have been compared to the manual dissection, which is considered as the reference model. These methods include a simple thresholding segmentation (Gonzalez and Woods (2002)), a thresholding with bone filling and skin detection, and the proposed partial volume approach. The first one can be considered as the base method, since only a simple thresholding segmentation is performed (first to second step of the pipeline represented in Figure 1, plus the sixth step to compute the LMP value). The second one is an extension of the base method, applying also a bone filling operation and the new skin detection step described in Section 2.2 (first to fourth step of the pipeline, plus the sixth step). Finally, the last method corresponds to

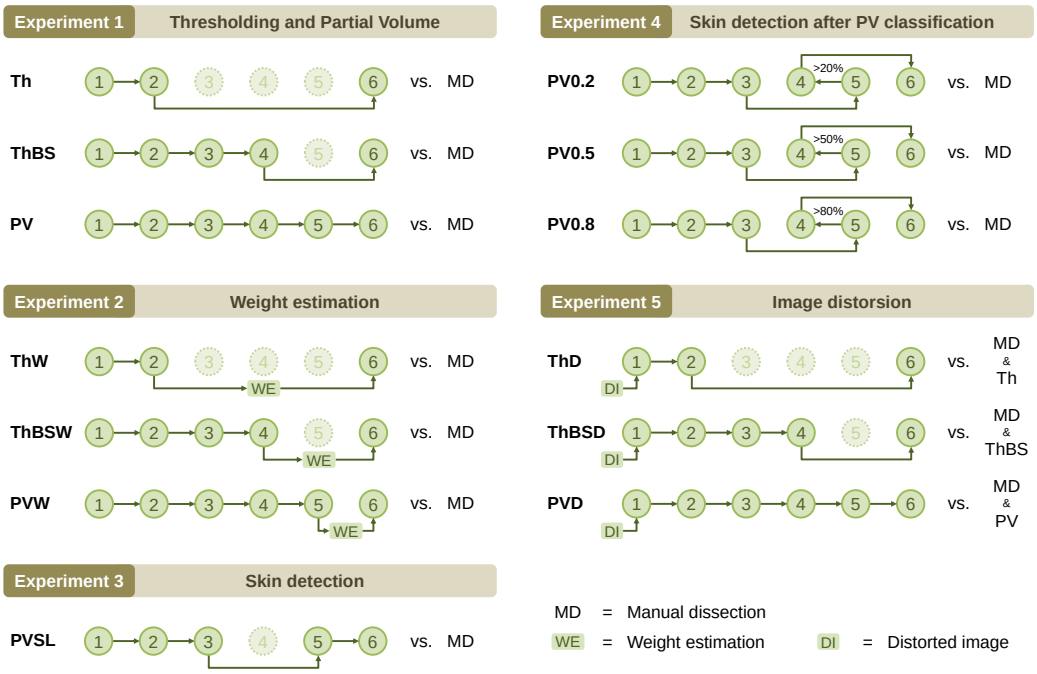


Figure 3: Steps from the whole pipeline which are part of the different alternatives analysed in each experiment.

the whole proposed approach, including the application of a partial volume model. The correlations of the different strategies have been computed to know which method better approximates the LMP values.

Experiment 2. While the CT scans of pig carcasses provide a good means to obtain the volume of each tissue by counting the number of voxels, the LMP values of the manual dissection are based on weights. To simulate this procedure, a density estimation is needed to compute the weight of the segmented tissues once the number of voxels of each one is known. However, the second experiment aims to show that the results of the algorithm are not improved when the density estimation is applied and the voxels are assigned

different weights, i.e. when the results take into account the weight, not the volume. As for the pipeline of this testing method, a new step corresponding to the weight estimation is included before the sixth step, which refers to the LMP computation. The model proposed by Vester-Christensen et al. (2009) has been followed to estimate the tissues' weights. The correlation between the weight-based results and the manual dissection has been computed, so that it can be compared with the results of the previous experiment.

Experiment 3. When using thresholds to segment the tissues from a carcass, the segmentation of the skin is troublesome, since the skin voxels have values very similar to the lean voxels' values. To solve this problem, all the voxels considered as lean which are at a distance of 3 mm from the background are considered as skin (see the fourth step of Figure 1). Thus, these voxels are not computed as lean when obtaining the LMP values, i.e. the skin tissue is not taken into account in the numerator of the LMP computation. The third experiment compares the results considering the skin as lean, i.e. ignoring the fourth step of the pipeline, with the results considering the skin as a new tissue. The correlation between the results considering the skin as lean and the manual dissection has been computed, so that it can be compared with the results of the first experiment, which considers the skin as a new tissue.

Experiment 4. The detection of the skin is based on thresholding methods, and it is performed before the partial volume classification. However, a voxel may be composed of skin and fat tissues at the same time (see Figure 2c), so the partial volume classification could be applied before the skin detection step. This is exactly the purpose of the fourth experiment, which swaps the

fourth and the fifth step of the pipeline. Since the partial volume classification outputs a tissue probability for each voxel, the probability threshold has to be defined in order to determine when a skin voxel can be considered as so. Assuming that the intensity values of skin voxels are similar to those of lean voxels, the lean tissue probability is taken into account. Hence, a voxel is considered to belong to the skin tissue when it is at 3 mm from the background and its probability to belong to the lean tissue is more than the defined threshold. Three different thresholds have been selected: 0.2 (20% probability of belonging to the lean tissue), since it is the value which obtains the best results; 0.5 (50% probability), since it is the more reasonable value to discern between fat and lean tissues; and 0.8 (80% probability), to include a more restrictive value. The correlation between the results swapping the fourth and fifth steps and the manual dissection has been computed, so that it can be compared with the results of the first experiment, which performs the skin detection step before the partial volume classification.

Experiment 5. Measurements of the same carcass using CT scanners from different vendors may show variation because of several factors, including the convolution kernel, reconstruction artefacts, beam hardening, spectral energy, and scatter, as well as variations in carcass size, shape, and position in scanner (Lamba et al. (2014); Mackin et al. (2015)). Moreover, although the difference is not so significant, measurements using the same scanner can also show variation (Jacobsen et al. (2016); Symons et al. (2016)). For this reason, the last experiment modifies the original CT images to simulate this image variability. Similarly to Bardera et al. (2014), a distortion function given by $\mathrm{HU}_{\mathrm{scale}} \times value + \mathrm{HU}_{\mathrm{shift}}$ has been applied to the values of all CT

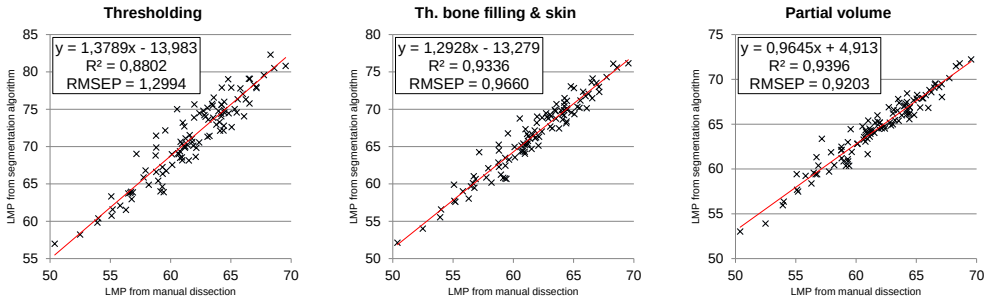
scan voxels, where HU_{scale} takes a value randomly generated between 0.97 and 1.03, and HU_{shift} takes a value between -20 and 20. Note that this distortion is different for each carcass, but the same for all voxels of the same carcass, so that noise is not added to the image, but only a global transform that will modify the histogram with a scaling factor and a shift. To find out which method best tolerates data variation, the correlation between each method with distortion and the manual dissection has been computed, as well as the correlation between the distorted results and the ones without distortion.

3. Results and discussion

In this section, the results of the experiments described in Section 2.5 are presented and discussed, always discerning between the partial dissection and the total dissection methodologies. The experiments aim to show how the proposed approach improves the LMP computation.

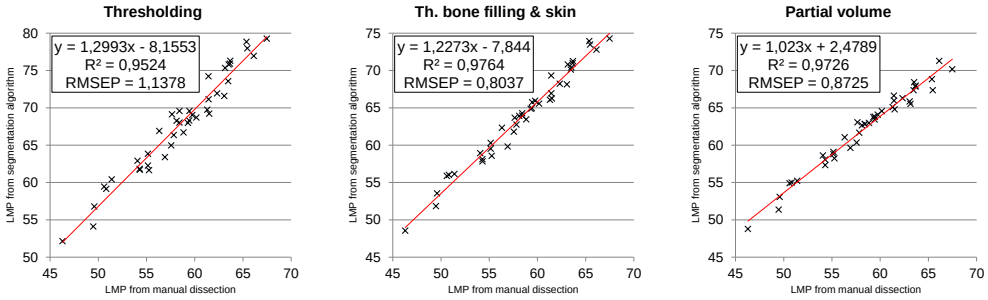
The results of the first experiment for the partial dissection methodology are shown in Figure 4, where the scatter plot between each method and the manual dissection is represented, and the R^2 and RMSEP values are given. The proposed approach (that is, the partial volume method) and the thresholding with bone filling and skin detection, which is an intermediate step of the proposed pipeline, clearly get the best results, with no significant differences between them (see Table 1 for the p-values). Similarly, the results for the total dissection methodology are shown in Figure 5, where there are no significant differences between the proposed approach and the thresholding with bone filling and skin detection either. However, although it has the lowest one, in this case the simple thresholding method also achieves a high

correlation, with nearly significant differences with respect to the proposed approach. As for the computational time, although the code has not been optimised, the bone filling and skin detection steps take most of the time of the two last compared methods. On the other hand, the simple thresholding and the partial volume classification need almost no time, so that the most efficient method is the simple thresholding (approximately 300 milliseconds), followed by the thresholding with bone filling and skin detection and the proposed approach, which take almost 5 seconds.



(a) (b) (c)

Figure 4: Correlation and error between each method and the manual dissection (partial dissection methodology).



(a) (b) (c)

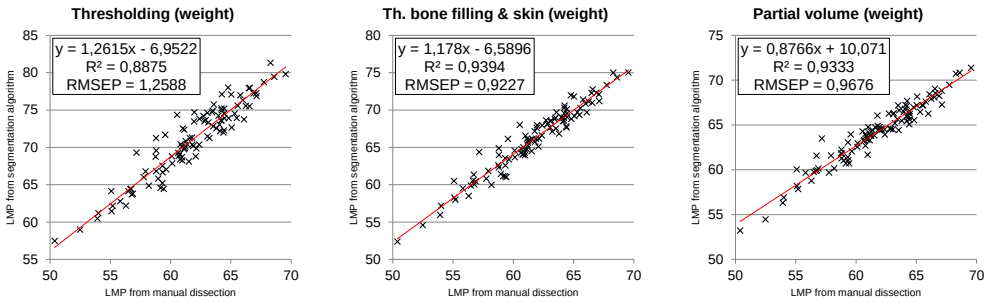
Figure 5: Correlation and error between each method and the manual dissection (total dissection methodology).

Dissection methodology	**Compared correlations**	**min**	**max**	**Steiger**
Partial	PV vs. Th	< 0.0001	0.0008	< 0.0001
	PV vs. ThBS	0.5161	0.5204	0.5195
Total	PV vs. Th	0.0597	0.0903	0.0616
	PV vs. ThBS	0.5978	0.6096	0.6072

Table 1: P-values of the comparison between the methods' correlations for the partial and the total dissection methodologies, showing the minimum and the maximum p-value from all the tests applied, and also the p-value of the Steiger's test (PV = Partial volume, Th = Thresholding, ThBS = Thresholding with bone filling and skin detection).

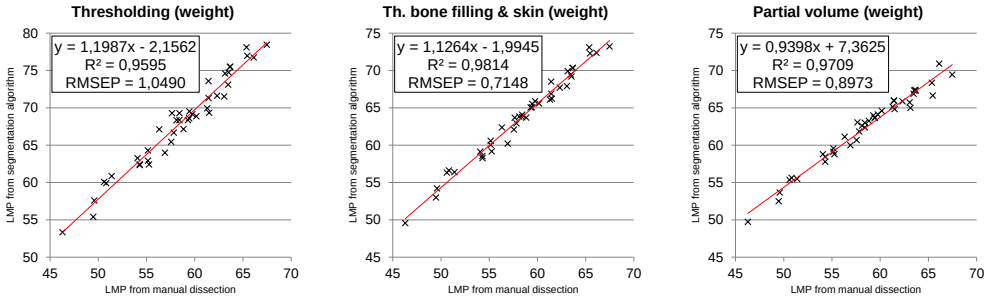
Regarding the second experiment, Figure 6 shows, for the partial dissection methodology and for each method, the scatter plot between the weight-based results (i.e. with density estimation) and the manual dissection, and also the R^2 and RMSEP values. Similarly, Figure 7 shows it for the total dissection methodology. For the thresholding-based methods, and interpreting

the results for both dissection methodologies, the correlation and the error lead to believe that the best option is to estimate the density, while for the proposed approach they suggest that the best option is to take into account only the volume, not the weight. However, and for the proposed approach, there are not significant differences for the total dissection methodology, and only some tests state that there are significant differences for the partial dissection methodology (see Table 2 for the p-values). Since the density estimation has not been able to obtain a better result for the proposed approach, it will not be added as a new step of the pipeline.



(a) (b) (c)

Figure 6: Correlation and error between each weight-based method and the manual dissection (partial dissection methodology).



(a) (b) (c)

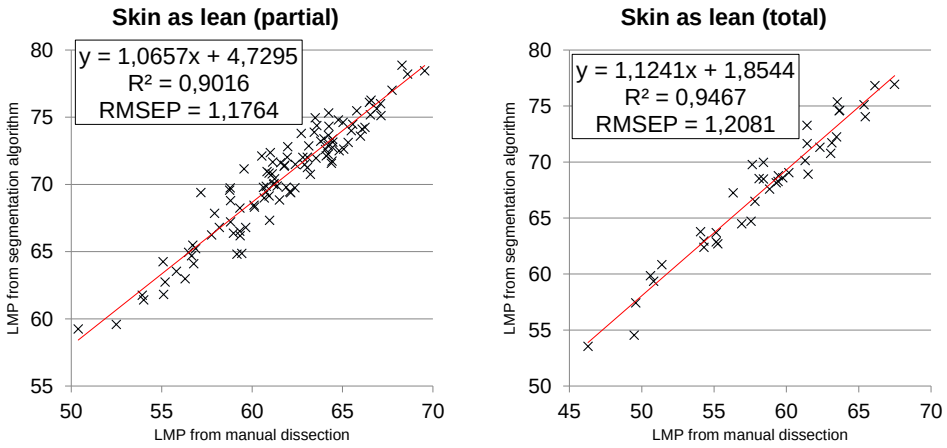
Figure 7: Correlation and error between each weight-based method and the manual dissection (total dissection methodology).

Dissection methodology	**Compared correlations**	**min**	**max**	**Steiger**
Partial	PVW vs. PV	0.0463	0.0589	0.0501
Total	PVW vs. PV	0.4930	0.5036	0.5019

Table 2: P-values of the comparison between the weight-based or volume-based correlations of the proposed approach for the partial and the total dissection methodologies, showing the minimum and the maximum p-value from all the tests applied, and also the p-value of the Steiger's test (PV = Partial volume based on volume, PVW = Partial volume based on weight).

The third experiment analyses the importance of classifying the skin as a new tissue in the fourth step of the proposed pipeline. For both dissection methodologies, Figure 8 shows the scatter plot between the results considering the skin as lean, i.e. ignoring the fourth step of the pipeline, and the manual dissection for the proposed approach, and also the R^2 and RMSEP values ($R^2 = 0.9016$ and RMSEP = 1.1764 for the partial dissection, and R^2

= 0.9467 and RMSEP = 1.2081 for the total dissection). The results considering the skin as a new tissue, and hence not considering it as lean when computing the LMP values, i.e. considering all the steps of the pipeline, are represented in Figure 4c and Figure 5c for the partial ($R^2 = 0.9396$ and RMSEP = 0.9203) and the total ($R^2 = 0.9726$ and RMSEP = 0.8725) dissection methodologies, respectively. Clearly, the latter are much better and have significant differences with respect to the former (see Table 3 for the p-values). Two reasons can explain this outcome. Firstly, the manual dissection takes into account the skin, so that it makes sense that detecting the skin in the segmentation step helps to improve the results. Secondly, some voxels which are close to the background may have a big uncertainty since it is difficult to know the tissue where they belong; when assigning some of these voxels to the skin tissue, the chances of assigning them to a wrong tissue disappear. Therefore, skin detection can be considered as a necessary step.



(a) (b)

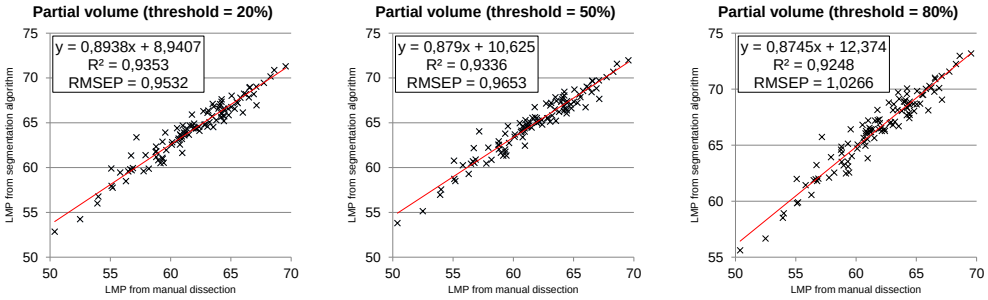
Figure 8: Correlation and error between the results of the proposed approach considering the skin as lean and the manual dissection (partial and total dissection methodologies).

Dissection methodology	**Compared correlations**	**min**	**max**	**Steiger**
Partial	PVSL vs. PV	< 0.0001	0.0009	0.0001
Total	PVSL vs. PV	0.0004	0.0141	0.0013

Table 3: P-values of the comparison between the proposed approach correlations considering the skin as lean (ignoring the skin detection step) or as a new tissue (considering all the steps) for the partial and the total dissection methodologies, showing the minimum and the maximum p-value from all the tests applied, and also the p-value of the Steiger's test (PV = Partial volume considering the skin as a new tissue, PVSL = Partial volume considering the skin as lean).

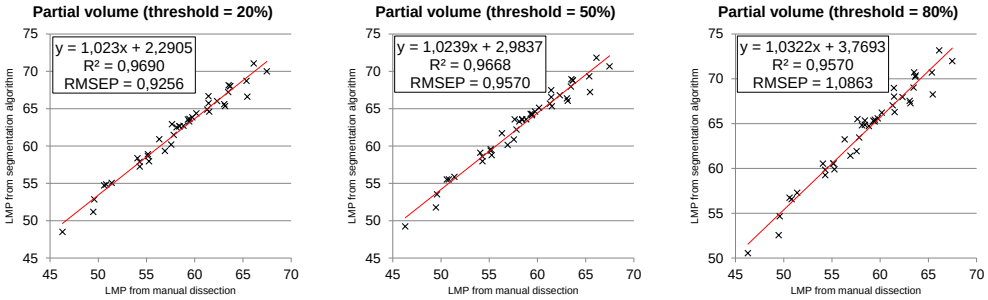
The fourth experiment evaluates the results when applying the skin detection step after the partial volume classification, i.e. swapping the fourth and the fifth step of the pipeline. In this way, every voxel has a certain prob-

ability of belonging to each tissue, so that a probability threshold is needed to determine if a voxel belongs to the skin tissue or not. Figure 9 shows, for the partial dissection methodology and for the proposed approach, the scatter plot between the results from the swapped pipeline and the manual dissection for each tested threshold, and also the R^2 and RMSEP values. Although the correlation is higher when using the original pipeline ($R^2 = 0.9396$ and RMSEP = 0.9203), only when the probability threshold is established to 0.8 ($R^2 = 0.9248$ and RMSEP = 1.0266) the differences are significant (see Table 4 for the p-values). Similarly, Figure 10 shows the same comparison for the total dissection methodology. In this case, the results obtained using the original pipeline ($R^2 = 0.9726$ and RMSEP = 0.8725) are significantly better than the ones using the swapped pipeline, regardless of the probability threshold used (0.2 threshold: $R^2 = 0.9690$ and RMSEP = 0.9256; 0.5 threshold: $R^2 = 0.9668$ and RMSEP = 0.9570; 0.8 threshold: $R^2 = 0.9570$ and RMSEP = 1.0863). Overall, placing the skin detection after the partial volume classification does not improve the results, and in some situations the results are worse. Hence, the original pipeline is preferred.



(a) (b) (c)

Figure 9: Correlation and error between the results of the proposed approach using the swapped pipeline (i.e. applying the skin detection step after the partial volume classification) and the manual dissection, taking into account three different probability thresholds, namely 0.2, 0.5 and 0.8 (partial dissection methodology).



(a) (b) (c)

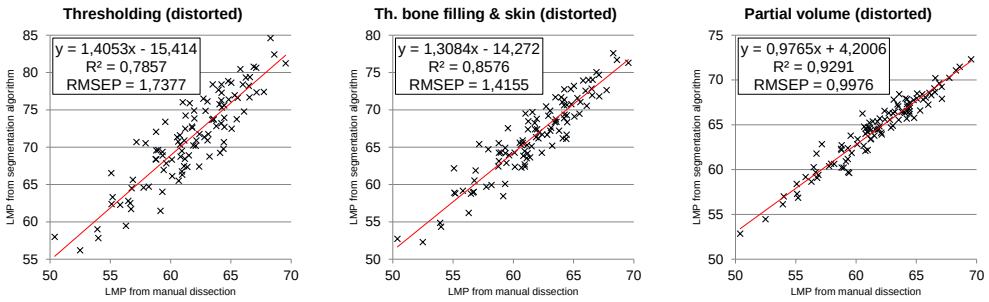
Figure 10: Correlation and error between the results of the proposed approach using the swapped pipeline (i.e. applying the skin detection step after the partial volume classification) and the manual dissection, taking into account three different probability thresholds, namely 0.2, 0.5 and 0.8 (total dissection methodology).

Dissection methodology	Compared correlations	min	max	Steiger
Partial	PV0.2 vs. PV	0.1925	0.1995	0.1957
	PV0.5 vs. PV	0.1130	0.1238	0.1167
	PV0.8 vs. PV	< 0.0001	0.0011	0.0001
Total	PV0.2 vs. PV	0.0029	0.0249	0.0070
	PV0.5 vs. PV	< 0.0001	0.0093	0.0005
	PV0.8 vs. PV	< 0.0001	0.0040	< 0.0001

Table 4: P-values of the comparison between the correlations of the proposed approach using the original pipeline or the swapped one (i.e. swapping the fourth and the fifth step of the pipeline) with three probability thresholds, namely 0.2, 0.5 and 0.8, for the partial and the total dissection methodologies, showing the minimum and the maximum p-value from all the tests applied, and also the p-value of the Steiger's test (PV = Partial volume; PV0.2, PV0.5 and PV0.8 = Partial volume with swapped pipeline and probability threshold established to 0.2, 0.5 and 0.8, respectively).

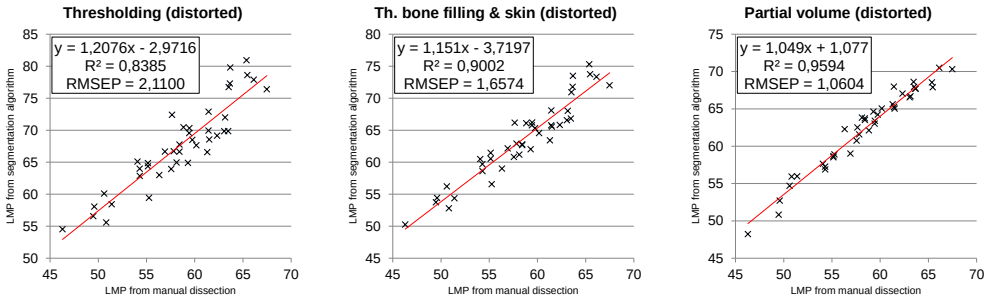
Taking into account the results from the first experiment, the proposed approach achieves results similar to the ones obtained using an intermediate step, i.e. the thresholding segmentation with bone filling and skin detection. However, one of the main goals of the proposed approach, which is analysed in the last experiment, is to be robust on data variability. Figure 11 shows the results for the partial dissection methodology, where the scatter plot between each method with distortion and the manual dissection is represented, and the R^2 and RMSEP values are given. Likewise, Figure 12 shows the results for the total dissection methodology. Furthermore, Figure 13 shows, for each method, the scatter plot between the distorted results and the ones without distortion, giving also the R^2 and RMSEP values. In this case,

all the carcasses are taken into account, since the values compared in the correlations are all obtained from the virtual methods, and no differentiation between manual dissection methodologies is needed. The results show that the proposed approach (the whole proposed pipeline) is the most robust method to image variability, obtaining correlation ratios significantly higher than the thresholding-based methods (see Table 5 for the p-values).



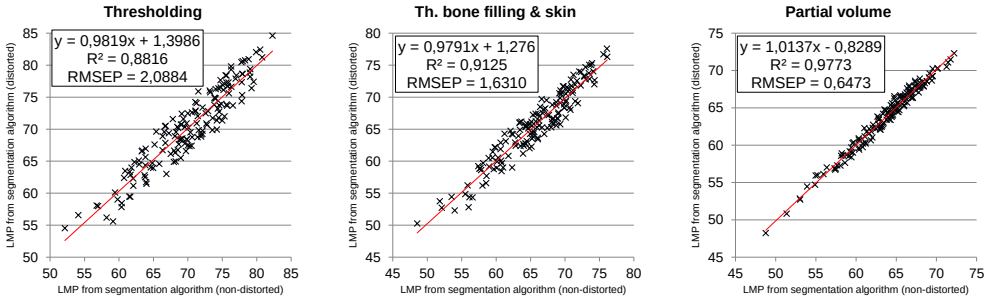
(a) (b) (c)

Figure 11: Correlation between each method with distortion and the manual dissection (partial dissection methodology).



(a) (b) (c)

Figure 12: Correlation between each method with distortion and the manual dissection (total dissection methodology).



(a) (b) (c)

Figure 13: Correlation between the distorted results and the ones without distortion for each method.

Dissection methodology	**Compared correlations**	**min**	**max**	**Steiger**
Partial	PVD vs. ThD	< 0.0001	< 0.0001	< 0.0001
	PVD vs. ThBSD	0.0003	0.0022	0.0004
Total	PVD vs. ThD	< 0.0001	0.0055	< 0.0001
	PVD vs. ThBSD	0.0029	0.0237	0.0045
None (dist. vs. non-dist.)	PVD vs. ThD	< 0.0001	< 0.0001	< 0.0001
	PVD vs. ThBSD	< 0.0001	< 0.0001	< 0.0001

Table 5: P-values of the comparison between the methods' correlations with distortion for the partial and the total dissection methodologies, and also between the methods' correlations with or without distortion, showing the minimum and the maximum p-value from all the tests applied, and also the p-value of the Steiger's test (PVD = Partial volume with distortion, ThD = Thresholding with distortion, ThBSD = Thresholding with bone filling and skin detection with distortion).

To summarise, five main conclusions can be drawn from these five experiments. The first experiment has shown that the results from the whole

proposed approach outperform the results from the simple thresholding, and that they are as acceptable as the ones obtained from a part of the same pipeline, i.e. the thresholding method with bone filling and skin detection, so they can both be used indistinctly. From the second experiment, the need of estimating the tissues' density could not be demonstrated, so that only the volume has been taken into account. The convenience to detect the animal skin has been evaluated in the third experiment, which has determined that the accuracy of the LMP computation is higher when considering the skin as a new tissue. The fourth experiment has evaluated the results of applying the skin detection after the partial volume classification, but there has been no evidence of improvement, so that the order of the steps has remained the same. Finally, the fifth experiment has tested the different methods with distorted images, and the results prove that the proposed approach is much more robust to data variability than the other thresholding-based methods.

Note that correlation is higher, and RMSEP is lower, when total dissection is considered instead of partial dissection. This is comprehensible since the carcasses were totally scanned, and this procedure is more similar to the total dissection than the partial dissection. In total dissection the lean of all the cuts of a carcass are separated manually and weighed; hence, the weight corresponds to the lean of the whole carcass (the same which has been scanned). In partial dissection, due to the reduction of the number of pieces to be dissected, the lean separated with a knife comes from the 4 main cuts and the tenderloin. Because of that, to obtain the LMP value of the whole carcass it is necessary to apply a scale factor, which was agreed to be the same for all EU countries (0.89) although there were some differences

between them. Thus, the use of this factor is a correction, and the LMP value of the whole carcass is estimated from the lean of the 4 main cuts plus the tenderloin. Furthermore, because of the way to compute the LMP values from 4 cuts, the cutting has an important effect on the obtained lean, and it is known that there are errors due to the cutting, especially in some cuts (Nissen et al. (2006)), that may affect the accuracy of the LMP prediction. The cutting errors are not so important in total dissection because all the cuts are dissected and the total lean is obtained by knife. Probably, scanning directly the main cuts would have given more precise results for partial dissection, because the scanned cuts would have been the same used in the prediction. In fact, Font-i-Furnols et al. (2009) showed a lower RMSEP for the prediction of lean meat content of the carcasses obtained by partial dissection when the four main cuts were scanned (0.71%) than when the whole carcass was scanned (0.82%).

4. Conclusion

In this paper, a six-step pipeline (carcass detection, image thresholding, bone filling, skin detection, partial volume classification and LMP computation), which includes a partial volume model and computes the LMP value of a pig carcass, has been evaluated. The method is based on an already presented pipeline, which has been extended by adding a new step to detect the animal skin in the thresholding stage. The method has also been thoroughly tested with 146 half carcasses, and compared with a simple thresholding method and a thresholding method with bone filling and skin detection, which corresponds to an intermediate step of the proposed pipeline (from the

first to the fourth step). Five experiments have been designed to evaluate the accuracy and robustness of the method as well as the necessity of every step of the pipeline. The results of these experiments determine that the proposed approach is an accurate method to compute the LMP values of pig carcasses from CT scans, and that it is not as affected by data variability as the other evaluated methods are.

In the future, we intend to improve the bone tissue model from the extended method, and apply the proposed approach to live pig CT scans. As for the latter, some efforts have been made to remove the internal organs which are present in the live pig CT images, but not required for the LMP computation (Xiberta et al. (2017)). Finally, the same automatic pipeline may be used to compute the LMP values of other species which may be of interest to the breeding companies and the meat industry.

Acknowledgement

This work has been funded in part by grants from the Spanish Government (Nr. TIN2016-75866-C3-3-R) and from the Catalan Government (Nr. 2017-SGR-1101), and has been carried out as part of the BR-UdG grant (research fellowship from the Universitat de Girona). CERCA Programme (Generalitat de Catalunya) is also acknowledged.

References

Anton Bardera, Rubén Martínez, Imma Boada, Maria Font-i-Furnols, and Marina Gispert. VisualPork: towards the simulation of a virtual butcher. In *FAIM I: First Annual Conference on Body and Carcass Evaluation,*

Meat Quality, Software and Traceability, pages 97–98, 2012. Held in Teagasc Food Research Centre, Ashtown, Dublin, Ireland, 24th - 26th September 2012.

Anton Bardera, Imma Boada, Albert Brun, Maria Font-i-Furnols, and Marina Gispert. Quantification of computed tomography pork carcass images. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1688–1692, 2014. doi: 10.1109/ICIP.2014.7025338.

Maren Bernau, Prisca Valerie Kremer, Elisabeth Lauterbach, Ernst Tholen, Brigitte Kirkegaard Petersen, Elke Pappenberger, and Armin Manfred Scholz. Evaluation of carcass composition of intact boars using linear measurements from performance testing, dissection, dual energy X-ray absorptiometry (DXA) and magnetic resonance imaging (MRI). *Meat Science*, 104:58–66, June 2015. ISSN 0309-1740. doi: 10.1016/j.meatsci.2015.01.011.

Gonca Bural, Drew Torigian, Sandip Basu, Mohamed Houseni, Ying Zhuge, Domenico Rubello, Jayaram Udupa, and Abass Alavi. Partial volume correction and image segmentation for accurate measurement of standardized uptake value of grey matter in the brain. *Nuclear Medicine Communications*, 36(12):1249–1252, December 2015. ISSN 0143-3636. doi: 10.1097/MNM.0000000000000394.

Anna Carabús, Marina Gispert, Albert Brun, Pedro López Rodríguez, and Maria Font-i-Furnols. In vivo computed tomography evaluation of the composition of the carcass and main cuts of growing pigs of three commercial

crossbreeds. *Livestock Science*, 170:181–192, 2014. ISSN 1871-1413. doi: 10.1016/j.livsci.2014.10.005.

Hwan Soo Choi, David R. Haynor, and Yongmin Kim. Partial volume tissue classification of multichannel magnetic resonance images - a mixel model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, September 1991. ISSN 0278-0062. doi: 10.1109/42.97590.

Emre Şener, Erkan U. Mumcuoglu, and Salih Hamcan. Bayesian segmentation of human facial tissue using 3D MR-CT information fusion, resolution enhancement and partial volume modelling. *Computer Methods and Programs in Biomedicine*, 124:31–44, February 2016. ISSN 0169-2607. doi: 10.1016/j.cmpb.2015.10.009.

Matthijs C. F. Cysouw, Gerbrand M. Kramer, Linda J. Schoonmade, Ronald Boellaard, Henrica C. W. de Vet, and Otto S. Hoekstra. Impact of partial-volume correction in oncological PET studies: a systematic review and meta-analysis. *European Journal of Nuclear Medicine and Molecular Imaging*, 44(12):2105–2116, November 2017. ISSN 1619-7089. doi: 10.1007/s00259-017-3775-4.

Gérard Daumas and Mathieu Monziols. An accurate and simple computed tomography approach for measuring the lean meat percentage of pig cuts. In *Proceedings of the 57th International Congress of Meat Science and Technology (ICoMST)*, pages 97–98, 2011. Held in Ghent, Belgium, 7th - 12th August 2011.

Birk Diedenhofen and Jochen Musch. cocor: a comprehensive solution for the

statistical comparison of correlations. *PLoS ONE*, 10(4):e0121945, April 2015. doi: 10.1371/journal.pone.0121945.

Andreas Dobrowolski, Róbert Romvári, Paul Allen, Wolfgang Branscheid, and Péter Horn. X-ray computed tomography as possible reference for the pig carcass evaluation [Schlachtkörperwertbestimmung beim schwein röntgen- computertomographie als mögliche referenzmethode]. *Fleischwirtschaft*, 84(3):109–112, 2004.

Juergen Dukart and Alessandro Bertolino. When structure affects function - the need for partial volume effect correction in functional and resting state magnetic resonance imaging studies. *PLoS ONE*, 9(12):e114227, December 2014. doi: 10.1371/journal.pone.0114227.

Maria Font-i-Furnols, Maria Fabiana Teran, and Marina Gispert. Estimation of lean meat content in pig carcasses using X-ray Computed Tomography and PLS regression. *Chemometrics and Intelligent Laboratory Systems*, 98 (1):31–37, 2009. ISSN 0169-7439. doi: 10.1016/j.chemolab.2009.04.009.

Maria Font-i-Furnols, Anna Carabús, Candido Pomar, and Marina Gispert. Estimation of carcass composition and cut composition from computed tomography images of live growing pigs of different genotypes. *Animal*, 9 (1):166–178, January 2015. doi: 10.1017/S1751731114002237.

L. E. Gangsei, J. Kongsro, E. V. Olsen, M. Røe, O. Alvseike, and S. Sæbø. Prediction precision for lean meat percentage in Norwegian pig carcasses using 'Hennessy grading probe 7': evaluation of methods emphasized at exploiting additional information from computed tomography. *Acta Agri-*

culturae Scandinavica, Section A — Animal Science, 66(1):17–24, May 2016. ISSN 0906-4702. doi: 10.1080/09064702.2016.1174292.

Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice-Hall, Upper Saddle River (NJ), USA, 2 edition, 2002. ISBN 0-201-18075-8.

Megan Jacobsen, Cayla Wood, and Dianna Cody. SU-G-206-07: dual-energy CT inter- and intra-scanner variability within one make and model. *Medical Physics*, 43(6Part25):3641–3641, June 2016. ISSN 2473-4209. doi: 10.1118/1.4956948.

I. Jansons, V. Strazdina, R. Anenkova, D. Pule, I. Skadule, and L. Melece. Development of new pig carcasses classification formulas and changes in the lean meat content in Latvian pig population. *Agronomy Research*, 14 (S2):1306–1314, 2016. ISSN 1406-894X.

Michael Judas, Reinbardt Höreth, and Wolfgang Branscheid. Computed tomography as a method to analyse the tissue composition of pig carcasses. *Fleischwirtschaft international*, 1/2007:56–59, 2007. ISSN 0179-2415.

Jørgen Kongsro, Morten Røe, Are Halvor Aastveit, Knut Kvaal, and Bjørg Egelandsdal. Virtual dissection of lamb carcasses using computer tomography (CT) and its correlation to manual dissection. *Journal of Food Engineering*, 88(1):86–93, September 2008. ISSN 0260-8774. doi: 10.1016/j.jfoodeng.2008.01.021.

P. V. Kremer, M. Förster, and A. M. Scholz. Use of magnetic resonance imaging to predict the body composition of pigs in vivo. *Animal: an*

562 *International Journal of Animal Bioscience*, 7(6):879–884, 2013. doi:
563 10.1017/S1751731112002340.

564 David H. Laidlaw, Kurt W. Fleischer, and Alan H. Barr. Partial-volume
565 Bayesian classification of material mixtures in MR volume data using voxel
566 histograms. *IEEE Transactions on Medical Imaging*, 17(1):74–86, February
567 1998. ISSN 0278-0062. doi: 10.1109/42.668696.

568 Ramit Lamba, John P. McGahan, Michael T. Corwin, Chin-Shang Li, Tien
569 Tran, J. Anthony Seibert, and John M. Boone. CT Hounsfield num-
570 bers of soft tissues on unenhanced abdominal CT scans: variability be-
571 tween two different manufacturers' MCDT scanners. *American Journal of*
572 *Roentgenology*, 203(5):1013–1020, November 2014. ISSN 0361-803X. doi:
573 10.2214/AJR.12.10037.

574 Sangdae Lee, Santosh Lohumi, Hyoun-Sub Lim, Takafumi Gotoh, Byoung-
575 Kwan Cho, and Samooel Jung. Determination of intramuscular fat con-
576 tent in beef using magnetic resonance imaging. *Journal of the Faculty*
577 *of Agriculture, Kyushu University*, 60(1):157–162, February 2015. ISSN
578 0023-6152.

579 Dariusz Lisiak, Kamil Duziński, Piotr Janiszewski, Karol Borzuta, and
580 Damian Knecht. A new simple method for estimating the pork car-
581 cass mass of primal cuts and lean meat content of the carcass. *Ani-*
582 *mal Production Science*, 55(8):1044–1050, 2015. ISSN 1836-0939. doi:
583 10.1071/AN13534.

584 Dennis Mackin, Xenia Fave, Lifei Zhang, David Fried, Jinzhong Yang, Brian

Taylor, Edgardo Rodriguez-Rivera, Cristina Dodge, Aaron Kyle Jones, and Laurence Court. Measuring computed tomography scanner variability of radiomics features. *Investigative Radiology*, 50(11):757–765, November 2015. ISSN 0020-9996. doi: 10.1097/RLI.0000000000000180.

Pia Marlene Nissen, Hans Busk, Marjatta Oksama, Marc Seynaeve, Marina Gispert, Pieter Walstra, Ingemar Hansson, and Eli Vibeke Olsen. The estimated accuracy of the EU reference dissection method for pig carcass classification. *Meat Science*, 73(1):22–28, May 2006. ISSN 0309-1740. doi: 10.1016/j.meatsci.2005.10.009.

Lucien Nocera and James C. Gee. Robust partial-volume tissue classification of cerebral MRI scans. In *Proceedings of SPIE 3034, Medical Imaging 1997: Image Processing*, 1997. doi: 10.1117/12.274118. Held in Newport Beach, CA, United States, 22nd - 28th February 1997.

Eli Vibeke Olsen, Lars Bager Christensen, and Dennis Brandborg Nielsen. A review of computed tomography and manual dissection for calibration of devices for pig carcass classification - evaluation of uncertainty. *Meat Science*, 123:35–44, January 2017. ISSN 0309-1740. doi: 10.1016/j.meatsci.2016.08.013.

Dzung L. Pham and Jerry L. Prince. Unsupervised partial volume estimation in single-channel image data. In *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. MMBIA-2000*, pages 170–177, 2000. doi: 10.1109/MMBIA.2000.852375. Held in Hilton Head Island, South Carolina, USA, 11th - 12th June 2000.

Róbert Romvári, Andreas Dobrowolski, Imre Repa, Paul Allen, Eli Vibeke Olsen, András Szabó, and Péter Horn. Development of a computed tomographic calibration method for the determination of lean meat content in pig carcasses. *Acta Veterinaria Hungarica*, 54(1):1–10, January 2006. ISSN 0236-6290. doi: 10.1556/AVet.54.2006.1.1.

Su Ruan, Cyril Jaggi, Jinghao Xue, Jalal Fadili, and Daniel Bloyet. Brain tissue classification of magnetic resonance images using partial volume modeling. *IEEE Transactions on Medical Imaging*, 19(12):1179–1187, December 2000. ISSN 0278-0062. doi: 10.1109/42.897810.

Lauren E. Salminen, Thomas E. Conturo, Jacob D. Bolzenius, Ryan P. Cabeen, Erbil Akbudak, and Robert H. Paul. Reducing CSF partial volume effects to enhance diffusion tensor imaging metrics of brain microstructure. *Technology & Innovation*, 18(1):5–20, May 2016. ISSN 1949-8241. doi: 10.21300/18.1.2016.5.

Peter Santago and Howard Donald Gage. Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Transactions on Medical Imaging*, 12(3):566–574, September 1993. ISSN 0278-0062. doi: 10.1109/42.241885.

James H. Steiger. Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2):245–251, March 1980. ISSN 0033-2909. doi: 10.1037/0033-2909.87.2.245.

Rolf Symons, Justin Z. Morris, Colin O. Wu, Amir Pourmorteza, Mark A. Ahlman, João A. C. Lima, Marcus Y. Chen, Marissa Mallek, Veit Sandfort,

and David A. Bluemke. Coronary CT angiography: variability of CT scanners and readers in measurement of plaque volume. *Radiology*, 281(3):737–748, December 2016. ISSN 0033-8419. doi: 10.1148/radiol.2016161670.

The Commission of the European Communities. Commission Regulation (EC) No 1249/2008 of 10 December 2008 laying down detailed rules on the implementation of the Community scales for the classification of beef, pig and sheep carcass and the reporting of prices thereof. *Official Journal of the European Union*, L 337:3–30, December 2008.

The European Commission. Commission Delegated Regulation (EU) 2017/1182 of 20 April 2017 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards the Union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses and live animals. *Official Journal of the European Union*, L 171:74–99, April 2017.

Jussi Tohka. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: a review. *World Journal of Radiology*, 6(11):855–864, November 2014. ISSN 1949-8470. doi: 10.4329/wjr.v6.i11.855.

Koen Van Leemput, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. A unifying framework for partial volume segmentation of brain MR images. *IEEE Transactions on Medical Imaging*, 22(1):105–119, January 2003. ISSN 0278-0062. doi: 10.1109/TMI.2002.806587.

Martin Vester-Christensen, Søren G. H. Erbou, Mads F. Hansen, Eli V.

Olsen, Lars B. Christensen, Marchen Hviid, Bjarne K. Ersbøll, and Rasmus Larsen. Virtual dissection of pig carcasses. *Meat Science*, 81(4):699–704, 2009. ISSN 0309-1740. doi: 10.1016/j.meatsci.2008.11.015.

Pieter Walstra and Gerard S. M. Merkus. Procedure for assessment of the lean meat percentage as a consequence of the new EU reference dissection method in pig carcass classification: based on discussion in the EU Management Committee on Pig Meat and based on discussions with dissection experts during a meeting on 18-19, 1994 at Zeist, NL. Technical Report ID-DLO 96.014, DLO Institute for Animal Science and Health, Lelystad, Netherlands, March 1996.

Pau Xiberta, Imma Boada, Anton Bardera, and Maria Font-i-Furnols. A semi-automatic and an automatic segmentation algorithm to remove the internal organs from live pig CT images. *Computers and Electronics in Agriculture*, 140:290–302, August 2017. ISSN 0168-1699. doi: 10.1016/j.compag.2017.06.003.