TESI DOCTORAL

Optimització del procés industrial d'escorxadors utilitzant tècniques de modelització, tractament de dades i classificació automàtica

Gerard Masferrer Caralt

UNIVERSITAT DE VIC
UNIVERSITAT CENTRAL DE CATALUNYA

Escola de Doctorat

TESI DOCTORAL

Optimització del procés industrial d'escorxadors utilitzant tècniques de modelització, tractament de dades i classificació automàtica

Gerard Masferrer Caralt

Directors de tesi: Pere Martí Puig; Moisés Serra Serra
Tutor: Ricard Carreras Ubach
Programa de doctorat: Ciències Experimentals i Tecnologies

2019

UNIVERSITAT DE VIC
UNIVERSITAT CENTRAL DE CATALUNYA
Escola de Doctorat

Amb el suport del Pla de Doctorats Industrials de la Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya. DI2015011

Generalitat de Catalunya
Departament d'Empresa i Coneixement
Secretaria d'Universitats i Recerca

DOCTORATS
INDUSTRIALS

Resum

Els escorxadors de porcí disposen d’una gran quantitat de dades a conseqüència del volum de caps sacrificats diàriament. Aquesta informació inclou per una banda dades intrínseques dels animals (sexe, raça, i pes), i per altra dades provinents de sistemes de classificació de canals.

Aquesta tesi està formada per un conjunt de publicacions on s'explora l'ús de tècniques de reconeixement de patrons en escorxadors. Aquestes tècniques s’utilitzen amb l'objectiu d'optimitzar el sistema productiu a través de la classificació de les canals segons característiques del pernil. La tesi es centra en la classificació de pernil, ja que és una de les peces més valuoses en la industria càrnia.

Dels resultats obtinguts se’n conclou que l’ús de models de reconeixement de patrons en escorxadors pot ser una eina útil per millorar els sistemes de classificació de canals, concretament per classificar pernils segons l’espessor de greix subcutani en el pernil i segons el pes. A més a més, aquests mètodes podrien permetre reduir personal, reduir errors deguts a la fatiga dels operaris, optimitzar processos posteriors a les sales d’especejament i incrementar la velocitat de la cadena de l’escorxador.

Resumen

Los mataderos de porcino disponen de una gran cantidad de datos a consecuencia del volumen de cabezas sacrificadas diariamente. Esta información incluye por un lado datos intrínsecos de los animales (sexo, raza y peso) y por otro lado datos obtenidos en los mataderos provenientes de sistemas de clasificación de canales.

Esta tesis está formada por un conjunto de publicaciones donde se explora el uso de técnicas de reconocimiento de patrones en mataderos. Estas técnicas se utilizan con el objetivo de optimizar el sistema productivo a través de la clasificación de las canales según características del jamón. La tesis se centra en la clasificación de jamón ya que el jamón es una de las piezas más valiosas en la industria cárnica.

De los resultados obtenidos se concluye que el uso de modelos de reconocimiento de patrones en mataderos, pueden ser una herramienta útil para mejorar los sistemas de clasificación de canales, concretamente para clasificar jamones según espesor de grasa subcutánea en el jamón y según peso. Además, estos métodos podrían permitir reducir personal, reducir errores debidos a fatiga de los operarios, optimizar procesos posteriores a las salas de despiece e incrementar la velocidad de la cadena del matadero.

Abstract

Pig slaughterhouses have a large amount of data as a result of the volume of heads slaughtered daily. This information includes intrinsic animal data (sex, breed and weight) on the one hand, and data from carcass classification systems in slaughterhouses on the other.

This thesis consists of a set of papers exploring the use of pattern recognition techniques in slaughterhouses. These techniques are used with the aim of optimising the production system through the classification of carcasses. Specificaly carcasses were classified according to the characteristics of the ham. The present thesis focuses on the classification of ham since ham is one of the most valuable pieces in the pork industry.

From the obtained results, it can be concluded that the use of pattern recognition models can be a useful online tool in slaughterhouses to classify hams according to the thickness of the subcutaneous fat. In addition, this method could allow the replacement of an operator in the production line, saving personnel costs, allowing faster chain speeds and reducing errors due to the fatigue of the operator.

M'agradaria agrair el suport que he tingut a l'hora de realitzar aquesta tesi.

He de donar les gràcies als meus directors, Pere Martí i Moisès Serra, per orientar-me i guiar-me durant el doctorat i al llarg de la meva trajectòria acadèmica.

A Mafrica per l'ajuda, suport i confiança que han dipositat en mi, ja que m'han permès desenvolupar la tesi; i en especial a en Ricard Carreras per ajudar-me sempre i animar-me a seguir en els moments difícils. És un plaer poder treballar al teu costat.

A la gent de IRTA per l'ajuda i col·laboració en els projectes.

A AGAUR i al programa de *Doctorats Industrials* de la Generalitat de Catalunya, per permetre amb aquests ajuts realitzar una tesi doctoral en una empresa.

I per descomptat, a la meva família i amics, que m'han recolzat en tot moment. I especialment a la Míriam, per fer-me gaudir de la vida al seu costat.

Gràcies, gràcies, gràcies.

ÍNDEX

ÍNDEX DE FIGURES

ÍNDEX DE TAULES

TESI COM A COMPENDI DE TREBALLS PRÈVIAMENT PUBLICATS

La present tesi doctoral, d'acord amb l'informe corresponent, autoritzat pels directors de tesi i l'òrgan responsable del programa de doctorat, es presenta com un compendi de tres treballs prèviament publicats. Les referències completes dels articles que pertanyen al cos de la tesi són els següents:

- **G. Masferrer**, R. Carreras, M. Font-i-Furnols, M. Gispert, P. Marti-Puig, and M. Serra, On-line Ham Grading using pattern recognition models based on available data in commercial pig slaughterhouses, Meat Science. 143 (2018) 39–45. doi:10.1016/j.meatsci.2018.04.011
- **G. Masferrer**, R. Carreras, M. Font-i-Furnols, M. Gispert, M. Serra, and P. Marti-Puig. Automatic ham classification method based on support vector machine model increases accuracy and benefits compared to manual classification, Meat Science. 155 (2019) 1-7. doi:10.1016/j.meatsci.2019.04.018
- **G. Masferrer**, R. Carreras, P. Marti-Puig, and M. Serra, Sorting Hams Using Bagged Decision Trees in a Commercial Pig Slaughterhouse, *Frontiers in Artificial Intelligence and Applications: Artificial Intelligence Research and Development* (2018) vol.(308) 84-88. doi:10.3233/978-1-61499-918-8-84

Dins el període de doctoral s'ha participat en els següents congresos amb una presentació i un pòster:

G. Masferrer, P. Marti-Puig, M. Serra, and R. Carreras Pattern Recognition Applied to Improve Pig Slaughterhouses Processes *9th International Congress on Environmental Modelling and Software*

G. Masferrer, R. Carreras, P. Marti-Puig, and M. Serra, Sorting Hams Using Bagged Decision Trees in a Commercial Pig Slaughterhouse *21st International Conference of the Catalan Association for Artificial Intelligence (CCIA 2018)*

I. INTRODUCCIÓ

I.I. Antecedents

I.I.I. Sector Porcí

En els darrers anys s'ha produït un augment de producció de carn a nivell mundial, passant de 272 milions de tones produïdes el 2007 a 334 milions de tones el 2017. A més, sembla que la tendència a nivell mundial seguirà augmentant en els propers anys (FAOSTAT 2017) (Fig I.1).

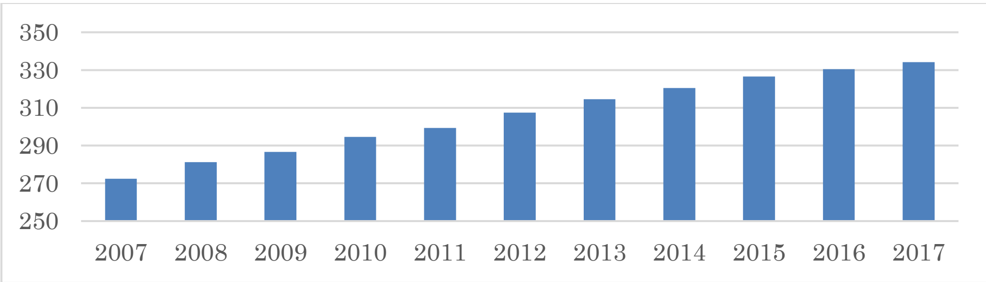

Fig. I.1 Milions de tones de carn produïdes anualment del 2007 el 2017.

Actualment la producció mundial de carn de porc és de 109,06 Milions de Tones (FAOSTAT 2017). Des d'un punt de vista d'espècies animals, la carn de porc és la més important i representa un 35,87% de la producció mundial, en segon lloc trobem la carn de pollastre amb un percentatge de 32,63%, seguida de la carn de boví amb un 19,82% i, finalment, altres espècies amb percentatges molt inferiors (Fig. I.2).

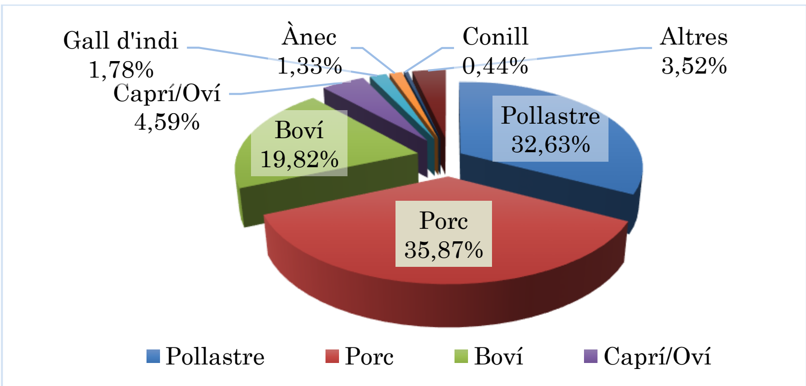

Fig. I.2 Percentatge de producció mundial segons tipus d'animal el 2017 (FAOSTAT).

Si observem la producció mundial des de un punt de vista regional, Àsia és el primer productor mundial de porc amb un 55,8% de la producció total: només a la Xina ja es produeixen més de 700 milions de porcs l'any. En segon lloc trobem Europa, que en produeix el 24,3% i on destaquen Alemanya i Espanya, amb 58,5 i 50,1 milions d'animals respectivament, seguits de França i Polònia amb 23,9 i 22,2 (Taula I). A Amèrica es produeix el 18,2% de la producció mundial amb EE.UU. situat com a principal productor amb 121,4 milions de porcs el 2017. Finalment trobem Àfrica, on es produeix un 1,2% del total i Oceania, amb un 0,5%.

Taula I Rànquing dels 10 països amb mes producció da caps de porc 2017 (FAOSTAT).

País	**Porcs produïts 2017**
1. Xina	702.021.000
2. Estats Units	121.390.200
3. Alemanya	58.408.370
4. Espanya	50.072.755
5. Vietnam	49.032.253
6. Brasil	43.185.385
7. Rússia	40.196.259
8. Filipines	27.142.373
9. França	23.858.700
10. Polònia	22.188.959

Observant aquests antecedents podem afirmar que el sector porcí espanyol és un dels més importants a nivell mundial. El 2018 el cens de porcs a l'estat espanyol va superar els 30 milions i representava el 20,8% de la producció de UE-28 (Eurostat 2018) (Fig I.3), tot i això, Alemanya el supera en producció anual (Taula I). Dins l'estat espanyol cal destacar Catalunya i Aragó: entre les dues representen el 52% del cens de porcí a l'estat amb un 26% cada una, seguides per Castella i Lleó i Andalusia amb un 13 i 9%, respectivament (MAPAMA 2017).

És importat destacar que a Espanya aproximadament un 10% de la producció de porc és porc ibèric. Aquesta producció està situada majoritàriament a Extremadura, on trobem aproximadament el 48% del cens d'ibèric, 1,2 milions de porcs al 2016 (MAPAMA 2017). Aquesta distribució és en gran part deguda a la producció espanyola de productes curats, destacant productes autòctons procedents del porc, entre ells una gran producció de pernil curat, tant procedent de porc ibèric com de porc blanc.

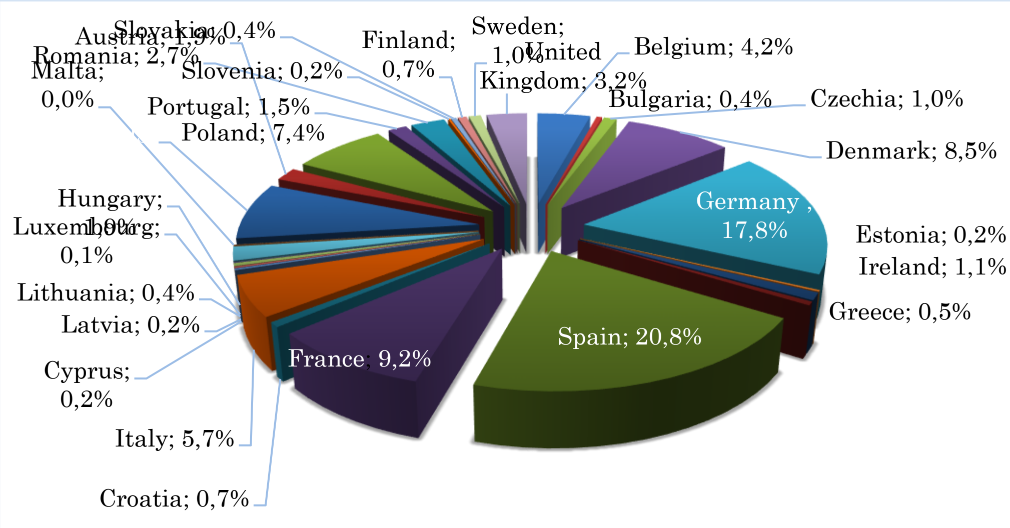

Fig. I.3 Percentatge de caps de bestià porcí a UE-28 el 2018 dades EURSTAT.

I.I.II. Importància del pernil en el sector porcí

El pernil és una de les parts més importants del porc. En primer lloc perquè representa entre un 25-30% de la canal del porc (Cisneros et al. 1996; Gispert et al., 2007), per altra banda perquè és una peça amb un alt contingut de carn magra i es tracta d'una peça molt apreciada per a la producció de productes curats i cuits.

Segons les característiques del pernil (pes, greix subcutani, percentatge de magre, raça del porc, greix intramuscular, perfil d'àcids grassos, contingut d'aigua lliure, etc.), s'utilitza per diverses finalitats, també influenciat per altres trets geogràfics i/o culturals (Dirinck, Van Opstaele, & Vandendriessche, 1997).

Per una banda atès el seu alt contingut de magre s'utilitza per produir carn fresca. El pernil es pot desossar i obtenir carn amb baix contingut de greix. Aquesta carn es pot filetejar, transformar en daus marinats per fer broquetes o moltes altres opcions. En aquest escenari, habitualment el pernil idoni seria un pernil que en el procés de desossar ens proporcionés un alt rendiment, pesos elevats, alt contingut de magre i mínim greix subcutani.

Per altra banda, un altre ús habitual és l'elaboració de pernil cuit, on la carn es sotmet a un procés d'injecció de salmorra seguit d'un procés de cocció en motlles. En aquest escenari el pernil idoni requereix habitualment una mica de greix per una bona cocció i un bon resultat organolèptic, però no un greix excessiu ja que en el procés de cocció es perdrà part d'aquest greix influint directament sobre el rendiment final. A més a més, una altra característica a tenir en compte és el pes, ja que al realitzar-se amb sistemes d'emmotllat habitualment es requereix un pernil fresc d'un pes determinat.

Finalment cal destacar també la producció de diverses varietats regionals de pernil curat a Europa. A Espanya s'elabora el pernil serrà o l'ibèric, molt apreciat per les seves característiques sensorials i organolèptiques, però no és l'única opció. A Itàlia trobem el prosciutto, amb una de les varietats més conegudes, el Prosciutto di Parma, elaborat amb pernils d'entre 7-10 kg, més grans que els serrans, amb una carn més rosada i amb uns gustos característics més suaus i dolços; o el Westphalian, a Alemanya, un pernil fumat i curat. En tots aquest processos de cocció i curació és molt important el pes i greix subcutani del pernil, ja que determinen directament la quantitat de sal i el temps de curació i/o cocció (Candek-Potokar & Skrlep, 2012).

En qualsevol d'aquests casos des de un punt de vista productiu és molt important la correcta caracterització, selecció i classificació del pernil segons les seves característiques.

I.II. Classificació de la canal en escorxadors

I.II.I. Percentatge de magre de la canal i sistema SEUROP

La Comunitat Econòmica Europea fa més de 50 anys que treballa en sistemes de classificació tècnica i avaluació de les canals de porc. En un primer moment es treballava des d'un punt de vista nacional/local, amb l'objectiu d'arribar a un model on els estats membres classifiquin les canals porcines de manera obligatòria segons un model comú basat en el percentatge de carn magra. L'objectiu final és el de fixar unes condicions comunes per l'intercanvi de canals porcines i la seva comparació. L'estimació del percentatge de magre ha de ser realitzat de manera objectiva.

Per tal de realitzar aquesta avaluació i classificació s'ha fixat un conjunt de mètodes basats en paràmetres tècnics de la canal i/o en relacions mètriques entre espessor de greix/magre en diversos punts de la canal (Font-i-Furnols et al., 2016). Per a realitzar aquestes mesures trobem mètodes manuals, semiautomàtics i automàtics, molts d'ells comuns a diversos països europeus. Els diferents mètodes estan regulats de manera global sota legislació europea comuna, però també regulats internament segons la legislació de cada país. Per tal d'aprovar un nou mètode en un país, cal realitzar un assaig nacional sota un control de la Comissió Europea supervisat per diferents experts dels estats membres. En aquests assaigs les canals utilitzades han de ser significativament representatives de la població porcina analitzada (de l'estat membre en qüestió) i que els resultats obtinguts siguin igualment significatius amb un marge d'error determinat.

Així doncs, segons el reglament delegat de la Comissió Europea (CE) 2017/1182, en els escorxadors és obligatòria la classificació de les canals segons el seu percentatge de magre i segons l'estàndard internacional SEUROP, que s'ha convertit en l'estàndard per a l'intercanvi de canals.

Per a la correcta classificació de la canal és necessària una presentació comuna de la canal, i el pesatge ha de realitzar-se amb l'animal en calent en un temps inferior a 45 minuts després del degollat.

Taula I-I. Valors de classificació SEUROP de canals porcines segons el seu percentatge de magre.

Percentatge de Magre (%)	Classe Comercial
Més de 60	S
Des de 55 fins a menys de 60	E
Des de 50 fins a menys de 55	U
Des de 45 fins a menys de 50	R
Des de 40 fins a menys de 45	O
Menys de 40	P

El sistema SEUROP estableix una categoria definida per una lletra segons el contingut de magre de la canal (Taula I-II). D'aquesta manera tots els escorxadors disposen de manera individualitzada de la classificació (percentatge i categoria) i del pes de cada una de les canals sacrificades.

Existeixen diversos mètodes autoritzats per a determinar el percentatge de magre. Per una banda trobem mètodes completament manuals com el mètode ZP (Zwei-Punkte Messverfahren), que es basa en la mesura amb una regleta de l'espessor de greix i de múscul en la canal. Aquest mètode només és vàlid en escorxadors petits amb un sacrifici inferior a 500 porcs la setmana. També trobem mètodes semiautomàtics basats en sondes de penetració de reflectància, com ara el Fat-O-Meat'er (FOM; Frontmatec Smørum A/S, Herlev, Dinamarca) o el Hennessy Grading Probe (HGP; Hennessy Grading System Ltd, Auckland, Nova Zelanda): aquests determinen el gruix del greix i de múscul en una posició anatòmica definida i els utilitzen per estimar el percentatge de magre.

Finalment trobem sistemes de classificació completament automàtics basats en diverses tecnologies: CSB Image-Meater o

VCS 2000, basats en sistemes d'anàlisi i processamet d'imatges (Font i Furnols & Gispert, 2009); gmSCAN, que està basat en la inducció electromagnètica, on s'analitza la variació d'un camp magnètic per determinar quantitat de carn (Fig. I.4); o Autofom III (Frontmatec Smørum A/S, Herlev, Denmark), que utilitza ultrasons per a determinar tant el percentatge de magre com les espessors de greix i magre en diversos punts de la canal.

Fig. I.4 Sistema gmScan escanejant porcs.

Molts d'aquets sistemes automàtics a part de proporcionar la informació sobre el percentatge de magre i classificació SEUROP poden proporcionar, en les versions més avançades, informació sobre el percentatge de magre dels principals talls primaris (pernil, ventre, espatlla, llom) així com estimacions del pes d'aquest. Addicionalment és possible disposar en alguns d'ells de la informació bruta procedent de la monitorització del porc.

I.II.II. Classificació del pernil

Habitualment en els escorxadors s'agrupen les canals segons les seves característiques. Les canals s'agrupen segons diferents criteris amb l'objectiu d'optimitzar processos productius, produir lots per a clients determinats, o optimitzar el "layout" de treballadors segons el producte produït.

Per a classificar les canals a la cambra d'oreig cada escorxador utilitza el seu propi criteri i aquest criteri pot variar molt segons l'escorxador, el país o el client final. A Espanya, a causa del mercat interior dedicat a la curació de pernil un dels criteris de classificació en les cambres d'oreig és segons les característiques del pernil.

El procés d'assignació de categoria habitualment es realitza a l'escorxador una vegada s'ha processat l'animal, s'ha pesat, sexat i s'ha escanejat per classificar (mitjançant sistemes semiautomàtics tipus FOM, o automàtics tipus AutoFoM III, gmScan, CSB Imatg Meater, etc).

Pel que fa a les característiques del pernil bàsicament intervenen tres criteris: pes, espessor de greix de cobertura i percentatge de magre.

I.II.II.I Classificació segons el pes del pernil

La classificació de la canal segons el pes del pernil es pot dur a terme de diverses maneres. El primer mètode i més simple és utilitzant el pes de la canal (pes en calent obtingut a l' escorxador) ja que el pernil representa entre un 25-30% de la canal. Tot i que utilitzant el pes de la canal ja s'obté una aproximació del pes del pernil, aquesta és millorable utilitzant el percentatge de magre per ajustar la classificació.

Per altra banda, si es disposa d'equips automàtics de classificació, aquests en ocasions poden predir el pes del pernil o és possible utilitzar la informació bruta dels equips per desenvolupar fòrmules de predicció pròpies de l'escorxador. En aquest context habitualment s'utilitzen arbres de decisió per crear grups basats en lots. Els grups s'estableixen segons rangs de pes (especificacions del client), un exemple podria ser el d'un grup per un client determinat que vol pernils arrodonits per fer curació, amb un rang de pes del producte acabat d'entre 12-14kg.

I.II.II.II. Classificació segons greix de cobertura

Dins el mercat nacional espanyol és habitual parlar de pernil gras, semi-gras, estàndard o prim. Aquestes categories fan referencia a la mínima espessor de greix observada en el *gluteus medius* una vegada el pernil està perfilat (arrodonit, polit i netejat amb una presentació com la que podem observar a Fig I.5 o Fig 2.1). Es tracta d'un valor categòric que tot i que no existeix un sistema de classificació estàndard amb un rang de greix perfectament definit, sí que permet establir un criteri aproximat com el que es presenta a la Taula I.III. Tot i això cada escorxador fixa el seu propi criteri segons especificacions del client, destí o processat del pernil.

Taula I-II Categoria de pernil segons greix subcutani de cobertura en el pernil.

Categoria de pernil	Espessor de greix (mm)
Prim	<10
Estàndard	entre < 10 i 15
Semigras	entre < 15 i 20
Gras	> 20

Cal destacar que aquesta classificació pot variar segons el processat final del pernil, de manera que si volem fer un pernil serrà (Fig I.5) amb una V necessitem lleugerament més greix per evitar clapes en el procés de treure la pell mecànicament.

Fig. I.5 Pernil acabat "serrà" amb forma de V.

Com s'ha comentat anteriorment no existeix un sistema estàndard establert per a fer aquesta classificació. En les sales d'especejament una vegada finalitzat el procés de perfilat dels pernils habitualment es classifiquen per inspecció visual. Tot i això, aquest pernils s'han classificat prèviament a l'escorxador per tal d'optimitzar la línia de producció i agrupar-los segons la categoria final resultant.

La classificació prèvia a l'escorxador, pel corresponent agrupament per categories a la cambra d'oreig, es pot fer mitjançant una classificació visual utilitzant un patró i un operari experimentat (mesurant el gruix de greix en el *gluteus medius* a la mitja canal en calent mitjançant un patró com el de la Fig I.6.)

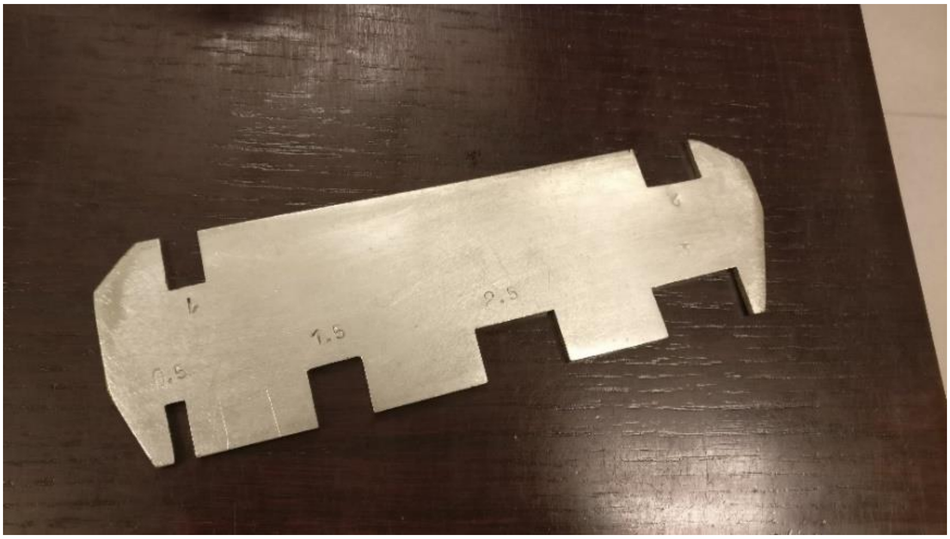

Fig. I.6 Patró per a la determinació de categories de pernil segons greix de cobertura.

O, de la mateixa manera que en el cas del pes del pernil sí es disposa d'equips automàtics de classificació, aquests poden proporcionar predictors de greix subcutani.

I.III. Dades disponibles en escorxadors

Avui en dia es sacrifica un gran volum de porcs que juntament amb la incorporació de noves tecnologies als escorxadors i a les sales d'especejament es genera una gran quantitat d'informació. Si seguim la cadena de producció del sector porcí s'observen diferents punts on es genera aquesta informació, a més moltes d'aquestes dades són de caràcter individual.

En primer lloc trobem dades procedents de la fase d'engreix. En aquest punt trobem dades de diferents tipologies: dades ambientals (p.ex.: temperatura i humitat relativa a les granges), alimentació (p.ex.: característiques dels pinsos, índex de conversió dels animals per fases, consums), pesos (en ocasions els sistemes de producció avançats ens permeten saber el pes dels animals diàriament o temporalment en diverses fases de l'engreix), espessors de greix (procedents d'ecografia en animals vius) i altres. Cal destacar que tot i l'existència d'aquestes dades a nivell de granja habitualment aquesta informació no arriba a l'escorxador, però sí que es realitza en situacions específiques, com estudis determinats, granges que són propietat de l'escorxador o que estan estretament lligades a aquest.

Un altre conjunt de dades disponibles correspon a les dades intrínseques dels animals. Aquí trobem dades de raça dels animals (genètica paterna i materna), sexe (mascle o femella), si s'ha castrat l'animal (ja sigui físicament o amb algun tractament d'inmunocastració) i el pes de la canal.

Finalment trobaríem multitud de dades procedents de sistemes de classificació de canals. La tipologia d'aquestes dades pot variar molt segons els equips disponibles i la tecnologia utilitzada. Destaca com a comuna i bàsica en tots els casos la informació del percentatge de magre de la canal i la seva corresponent classificació SEUROP, ja que es tracta d'una dada obligatòria.

I.III.I. Dades disponibles en els treballs publicats

A l'escorxador on s'ha realitzat aquesta tesi, Matadero Frigorífico del Cardoner SA (Mafrica), es sacrifiquen anualment uns 500.000 porcs amb una freqüència de 1600-2000 porcs diaris. En la Fig.I.7 es pot observar, de manera esquemàtica, la informació disponible en la base de dades de l'escorxador.

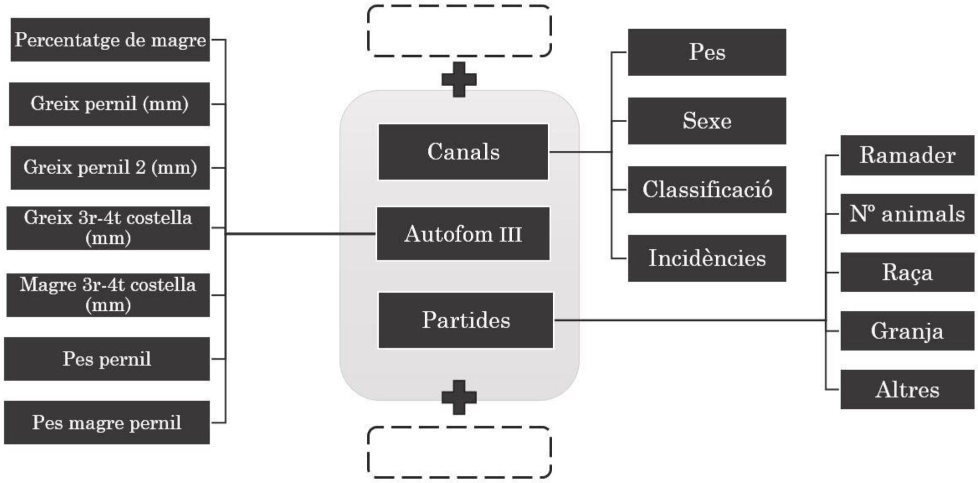

Fig. I.7 Esquema de les dades disponibles a l'escorxador.

A més de les dades descrites anteriorment es disposa d'un Autofom III (Frontmatec Smørum A/S, Herlev, Denmark). Aquest proporciona informació sobre el percentatge de magre i la classificació de la canal. Al tractar-se de la versió bàsica no es disposa dels paràmetres bruts, però tot i així es poden conèixer diferents paràmetres relacionats amb el pernil i la canal. Entre aquets paràmetres trobem informació sobre l'espessor de greix i magre entre la 3ra i 4ta costella (predicció del valor que obtindríem amb un sistema semiautomàtic FOM – Fat-O-Meat'er), l'espessor de greix mínima en dos punts del pernil (sobre el múscul *gluteus medius*, i en la part cranial d'aquest) i prediccions del pes del pernil (pes amb os, pes sense os i pes del magre).

Finalment es disposa de dades de classificació del pernil (classificació categòrica segons el greix de cobertura del pernil) obtingudes manualment per un operari qualificat. Aquesta informació correspon a l'espessor de greix mesurat a la mitja canal sobre el múscul *gluteus medius* utilitzant un patró (Fig. I.6) i informació relacionada amb tot tipus d'incidències que s'hagin pogut produir a l'escorxador.

Cal destacar que l'estructura d'informació en la present tesi s'ha anat incrementant i optimitzant gràcies als treballs duts a terme durant el doctorat i la incorporació de nous equips i tècniques.

Entre aquesta nova informació trobem la incorporació de gmSCAN ™ (gmStel i Lenz Instruments S.L) un nou equip per a la classificació de canals aprovat a Espanya per la Comissió Europea, l'assaig de validació del qual s'ha efectuat a Mafrica durant el període de doctorat. Aquest equip utilitza la inducció magnètica (variació d'un camp magnètic al passar la canal pel mig) per a predir la quantitat de carn de la canal i juntament amb la informació de pes, predir el percentatge de magre de la canal, el percentatge de magre dels principals talls primaris i predir també el pes dels principals talls.

A més a més, en un futur, s'incorporarà informació sobre el percentatge de greix intramuscular en canals, com a resultat del projecte "Millora integral de la carn de porc a través d'estratègies productives a nivell de granja i eines innovadores, NIRS (Near Infrared Spectroscopy) on-line per la seva classificació, a nivell d'escorxador. 562104820163A" també dut a terme durant el període de doctorat industrial. És més que probable que aquesta informació augmenti en els propers anys.

I.IV. Treball previ a l'escorxador

Per al correcte acompliment d'aquest treball i la conseqüent publicació dels articles que constitueixen aquesta tesi s'ha hagut de dur a terme un conjunt de tasques. Aquesta feina prèvia, tot i no formar part dels treballs publicats, resulta indispensabla per el correcte desenvolupament del doctorat i contribueixen a defensar la hipòtesi principal de la tesi (apartat II).

Inicialment tot el treball va començar amb la introducció i comprensió del sistema de producció, funcionament de l'escorxador i sala d'especejament per la posterior comprensió de la informació generada i el seu processament.

I.IV.I. Traçabilitat de canals amb RFID

En tot procés és important la integritat de les dades però si volem utilitzar-les amb tècniques de reconeixement de patrons i processat encara resulta més important partir d'un bon sistema d'informació. En el cas de l'escorxador la traçabilitat de la canal es realitza mitjançant un sistema de tags RFID (Radio Frequency Identifier). Després del procés de degollat i preescaldat, les canals de porc son penjades verticalment en la cadena de treball de l'escorxador. En aquest moment a cada un dels peus del porc se li posa un ganxo on hi ha integrat un tag RFID (Fig.I.8). D'aquesta manera cada animal va identificat amb dos tags, cadascun corresponent a cada mitja canal.

Fig. I.8 Canal de porc amb dos ganxos passant per una antena de lectura RFID.

Al llarg de tot el recorregut per l'escorxador hi ha antenes en cada un dels punts on s'introdueix nova informació de la canal o es llegeix aquesta per la distribució i classificació.

Entre els punts on s'introdueix informació trobem el sistema inicial d'assignació de ID (punt on s'assigna un identificador únic a cada porc relacionat amb els dos tags que utilitza), Autofom III (escàner del porc i emmagatzematge de tots els predictors i percentatge de magre), gmScan (escàner del porc i emmagatzematge del percentatge de magre i predicció del pes dels principals talls primaris), punt d'inspecció veterinària/punt de control crític (emmagatzematge de tot tipus d'incidències de la canal), bàscula (lectura de tots els paràmetres anteriors i emmagatzematge del pes i la classificació de la canal) i finalment el sistema de distribució de les canals per vies compostes de diversos punts (per agrupar i distribuir les canals segons les característiques).

Cada un dels punts requereix d'una correcta lectura dels dos tags per emmagatzemar correctament la informació i cada un d'ells és susceptible d'introduir errors. A causa d'això inicialment hi havia un percentatge superior al 40% d'animals en què es detectava algun error en l'emmagatzematge de la informació. Els principals problemes detectats eren els següents:

- Problemes amb l'orientació d'antenes amb ratis d'errors de lectura significatius.
- Problemes amb la velocitat de la línia i punts on s'havien de llegir dos xips situats un al costat de l'altre.
- Identificador ID de la canal que contenien tags de mitja canal d'un porc i mitja canal de la següent.
- Percentatge significatiu de xips defectuosos degut al desgast en la utilització habitual.

Per aquests problemes diversos es produïen bastants errors de classificació del pernil degut a males assignacions d'informació (dades corresponents al porc anterior o posterior, mitges canals no classificades per xips illegibles, entre altres). A més a més, el conjunt d'errors impossibilitava el correcte anàlisi de dades per l'avaluació de partides de proveïdors.

És per això que inicialment es van dur a terme diferents treballs previs per a maximitzar la qualitat de la informació. Per una banda, es van analitzar cada un dels punts de lectura, es va estudiar l'orientació de les antenes, i es va optimitzar la orientació del camp magnètic generat per aquestes per tal de maximitzar les respostes dels tags.

En el primer punt de la cadena, on s'assigna l'identificador únic a la canal i els corresponents dos xips, es va instal·lar un sistema amb doble antena de lectura. En aquest punt (Fig. I.8) la canal passa de manera horitzontal amb dos tags en paral·lel, com a conseqüència en un percentatge elevat de canals només es llegia un xip ja que amb el sistema RFID es llegia múltiples vegades un tag i no es podia llegir el del costat. D'aquesta manera aseguràvem l'assignació de tags inicial correcta pràcticament en la totalitat de les canals i disposàvem d'un sistema de backup en cas d'averia.

Dins aquest conjunt de tasques també cal destacar petits algoritmes desenvolupats per detectar desfasament de les dades i per quantificar xips defectuosos per la seva eliminació (degut als múltiples punts de lectura en l'escorxador ens permetia analitzar les seqüencies de xips de lectura i creuar les dades dels múltiples punts per detectar errors d'assignacions i xips defectuosos).

Com a resultat de tots aquests treballs s'ha aconseguit una millora substancial en les dades i en els sistemes de classificació, millorant la classificació de les canals i/o pernils tot i que no s'ha quantificat. Tot i això en els diferents treballs publicats en el marc de la tesi, com a norma general, es netejaven i eliminaven les mostres que presentaven algun error ja que es disposava d'un gran volum de dades.

II. OBJECTIUS

II.I. Hipòtesi de treball

Dels antecedents exposats als apartats anteriors, tenint en compte el funcionament dels escorxadors i de les sales d'especejament i la quantitat de les dades obtingudes s'ha formulat la següent hipòtesis:

> *"L'aplicació de tècniques de processat de dades i reconeixement de patrons, utilitzant dades procedents del procés productiu i noves tecnologies, millora el rendiment econòmic i productiu d'escorxadors i sales d'especejament de porcí."*

II.II. Objectius segons articles

El següent treball està format per tres articles publicats en revistes i congressos, cada un d'ells amb uns objectius concrets que s'exposaran a continuació. Tots ells s'emmarquen en un context general de classificació de pernils en escorxadors. El treball realitzat s'emmarca dins d'un ajut de Doctorat Industrial, pel que les publicacions científiques s'enfoquen des d'un punt de vista de necessitat de la indústria.

II.II.I. Comparació de diversos mètodes de reconeixement de patrons per la classificació de pernils

A Mafrica, l'escorxador on s'ha realitzat aquest estudi, es classifiquen les canals segons dos criteris principals, pes del pernil i greix subcutani del pernil.

En l'inici del projecte la classificació segons greix de cobertura en el pernil es duia a terme de forma completament manual. Un operari assignava una categoria (gras, semigras, estàndard o prim) en funció del greix observat en el pernil. A més es disposava de molta informació individualitzada dels porcs sacrificats, informació procedent de sistemes automàtics de classificació de canals (Autofom III) i informació de raça, sexe i pes.

Els objectius de l’estudi es divideixen en tres parts:

- Classificar els pernils en base a la predicció de l’espessor del greix subcutani dels pernils mitjançant la aplicació i avaluació de diferents tècniques (arbres de decisió, KNN, SVM, LDA/nLDA) . Com a dades d’entrada s’utilitzen les dades proporcionades per l’Autofom III i les dades intrínseques de l'animal i la resposta a imitar és la classificació realitzada manualment per un operari.
- Avaluar l'impacte de cada predictor en la precisió de la classificació dels pernils.
- Avaluar diverses combinacions de predictors disponibles en diferents escenaris d'escorxadors i comparar-los.

II.II.II. Comprovació de resultats entre un mètode de classificació basat en SVM i una classificació manual

Del resultat del primer estudi realitzat, s’han obtingut diversos models que prediuen la classificació del pernil que realitzaria l’operari. Alguns d’aquests models s’han implementat a l’escorxador per fer la classificació de manera automàtica.

Els objectius d’aquest article són:

- Comparar el millor model obtingut en el primer article, amb la classificació realitzada manualment per un operari.
- Avaluar l'efecte econòmic de la precisió del mètode de classificació del pernil per a la indústria càrnia.

II.II.III. Millorar l’agrupació delots de pernil segons pes utilitzant arbres de decisió i informació complementaria de l’escorxador

En el funcionament habitual de la sala d’especejament s’acostuma a agrupar els pernils en lots segons especificacions de clients. D’entre les especificacions dels clients un dels criteris utilitzats és el rang de pes del pernil resultant. Per a classificar els pernils a l'escorxador es poden utilitzar arbres de decisió simples basats en el pes en canal del porc i el seu percentatge de magre.

L'objectiu d'aquesta publicació és comparar tres models basats en arbres de decisió per a classificar pernils segons categories de pes (rangs de pes preestablerts).

- Arbre de decisió utilitzant només el pes de la canal i el percentatge de magre.
- Arbre de decisió utilitzant com a entrada el pes de la canal, el percentatge de magre, el sexe, la raça i la classificació de pernil segons el greix subcutani.
- Model amb arbres embossats "*Bagged trees*" utilitzant com a entrada el pes de la canal, el percentatge de magre, el sexe, la raça i la classificació de pernil segons el greix subcutani.

1. ONLINE HAM GRADING USING PATTERN RECOGNITION MODELS BASED ON AVAILABLE DATA IN COMMERCIAL PIG SLAUGHTERHOUSES

On-line Ham Grading using pattern recognition models based on available data in commercial pig slaughterhouses

Gerard Masferrer[ab], Ricard Carreras[b], Maria Font-i-Furnols[c], Marina Gispert[c], Pere Marti-Puig[a], Moises Serra[a]

[a]*Information and Digital Technology Department, UVic-UCC, Vic (Barcelona)*

[b]*Mafrica.SA, Paratge Can Canals Nou, S/N 08250 Sant Joan de Vilatorrada (BCN)*

[c]*IRTA-Product Quality, 17121 Monells, Girona, Spain*

Abstract

The thickness of the subcutaneous fat in hams is one of the most important factors for the dry-curing process and largely determines its final quality. This parameter is usually measured in slaughterhouses by a manual metrical measure to classify hams. The aim of the present study was to propose an automatic classification method based on data obtained from a carcass automatic classification equipment (AutoFom) and intrinsic data of the pigs (sex, breed, and weight) to simulate the manual classification system. The evaluated classification algorithms were decision tree, support vector machines (SVM), k-nearest neighbour and discriminant analysis. A total of 4000 hams selected by breed and sex were classified as thin (0-10mm), standard (11-15 mm), semi-fat (16-20 mm) and fat (>20 mm). The most reliable model, with a percentage of success of 73%, was SVM with Gaussian kernel, including all data available. These results suggest that the proposed classification method can be a useful online tool in slaughterhouses to classify hams.

Keywords dry-cured hams; ham-fat grading; subcutaneous fat thickness; pattern recognition

1.1. Introduction

Ham is one of the most valued product in pork meat industry. This primal cut represents between 25 and 30 percent of the carcass (Cisneros, Ellis, & McKeith, 1996; Gispert et al., 2007) and is the basis of different regional specialities focused on preserving and flavouring raw meat (Dirinck, Van Opstaele, & Vandendriessche, 1997). Those specialities include different techniques such as salting dry-cured ham, smoking or wet curing. Some examples are Westphalian ham in Germany, Prosciutto in Italy, and Jamon Serrano in Spain.

The Subcutaneous Fat Thickness (SFT) in hams determines, among other factors, which is the best process for the ham to be submitted. Hams with low subcutaneous fat have a high lean meat percentage (LMP) and are more appropriate to be processed as raw or cooked meat while hams with higher subcutaneous fat are more appropriate to be cured or smoked.

Moreover, the SFT determines the optimum curing time (Bosi, Russo, & Paolo, 2004), which is directly related to the quality of the final product (Čandek-Potokar & Škrlep, 2012). Therefore, classify the ham according to the SFT is crucial to get the maximum benefit of the product, in economic and quality terms.

The thickness of the subcutaneous fat is determined by several factors, among which can be highlighted the breed (Gispert et al., 2007; Wood et al., 2004), the sex (Font-i-Furnols et al., 2012; Gispert et al., 2010), the slaughter weight (Fàbrega et al., 2011; Latorre, García-Belenguer, & Ariño, 2008) and the diet (Realini et al., 2010; Tous et al., 2014; Wood et al., 2004). Regarding the breed, there are leaner breeds, as would be the Pietrain and other fattier breeds such as the Duroc (Cilla et al., 2006; Edwards, Bates, & Osburn, 2003). In terms of sex, females tend to deposit more subcutaneous fat than males (Gispert et al., 2010; Wood, Enser, Whittington, Moncrieff, & Kempster, 1989). Moreover, the castration, especially surgical but also immunological, also contributes to deposit more subcutaneous-fat compared with entire male pigs (Gispert et al., 2010; Wood et al., 2008).

Nowadays slaughterhouses have different methods to estimate the SFT of hams. One of the most used method is the visual system based on a metrical measure of the SFT over the *Gluteus medius* muscle, similar to ZP (Zwei-Punkte Messverfahren) measures, used to determine carcass LMP (Daumas, 2011; Font-i-Furnols et al., 2016). Indeed, the carcass LMP is a parameter widely used in slaughterhouses as the current EU legislation establishes it as compulsory for carcasses classification. There are different methods to determine LMP based, predominantly, on the existing relationship of thickness between fat and muscle in several parts of the carcass (Font i Furnols & Gispert, 2009).

Obtaining these measures manually is unsuitable in slaughter plants with medium/high speed line, therefore the most used methods to determine LMP are semiautomatic systems based on reflectance penetration probes, as for instance the Fat-O-Meat'er (FOM; Frontmatec Smørum A/S, Herlev, Denmark) or the Hennessy Grading Probe (HGP; Hennessy Grading System Ltd., Auckland, New Zealand), which determine fat and muscle thickness at a defined anatomical position and use them to estimate carcass LMP. Alternatively, there are non-invasive and fully automatic systems such as AutoFom (Frontmatec Smørum A/S, Herlev, Denmark) which is based on three-dimensional ultrasonic systems, or VCS 2000 (e + V Technology GmbH, Oranienburg, Germany) that extracts LMP by processing and analysing images (Font i Furnols & Gispert, 2009). Some of these devices also can estimate several SFT at the loin and at the ham level. For instance, AutoFom, provides several SFT parameters of the ham like *fatham2* (minimum subcutaneous fat plus skin thickness measured with a ruler over the muscle *Gluteus medius*) and *fatham3* (thickness of the subcutaneous fat plus skin measured with a ruler, perpendicularly to the skin, at the cranial part of muscle *Gluteus medius*).

Other systems, such as thermography technology have been proposed to classify the hams according to the SFT, being the hams with lower fat cover the ones that display a significantly warmer average temperature surface (Nanni Costa et al., 2010). Also computed tomography has been used in experimental conditions to determine the fat thickness at different anatomical positions mainly in the loin region (Lucas et al., 2017) although it could also been used in the ham region as has been done in live pigs (Carabús et al., 2014).

Nowadays a certain amount of data is collected in the slaughter line like gender and carcass weight, but also much other information from the productive chain is available such as breed, diet, transport and farm conditions, medication and castration (if done). In this context, with all this available data it is possible to take technical and commercial real-time decisions to better classify products and maximise profits. Therefore, our hypothesis is that complementing the Autofom-III set of estimated parameters with those additional ones could be used to improve the ham classification rate according to the SFT.

To carry out this classification it is possible to use classifiers. A classifier is an algorithm used to assign an unlabelled incoming element in a known category based on certain characteristic information of that element. These algorithms need to perform a learning stage. There are two types of primary learning strategies: supervised learning which elaborates a mathematical function (hypothesis) from previously labelled training data and unsupervised learning which does not have a training package that allows knowing the data labels, so it is necessary to use grouping techniques that try to build these labels. Among supervised algorithms, some of the most widespread are *Decision Trees*, *K-Nearest Neighbours (KNN), Linear and Nonlinear Discriminant*

Analysis (LDA/nLDA) and Support Vector Machine (SVM) (Bishop, 2006). Between the unsupervised classifiers the most popular strategies are the clustering which includes the Hierarchical and k-Means clustering algorithms.

Thus, the objectives of this study are: (1) To apply and assess different supervised classification techniques (Decision trees, kNN, SVN, LDA/nLDA) to predict the classification of hams according to SFT by combining data form Autofom III and intrinsic data from the animal, (2) to evaluate the impact of each predictor in the accuracy of ham classification, and (3) to evaluate several combinations of predictors available in different slaughterhouses scenarios and to compare them.

1.2. Material and Methods

1.2.1. Animals and facilities. The dataset construction

This study was carried out with data obtained during May 2016 from pigs fattened in Spanish commercial farms and slaughtered in a commercial slaughterhouse (MAFRICA S.A.) located in Sant Joan de Vilatorrada, Catalonia, Spain. All farms were less than 200 km far from the slaughterhouse and pigs were transported using trucks in groups (usually of between 80 and 220 animals). Once in the slaughterhouse pigs rested into lairage pens between 2 and 4 hours before being slaughtered.

This slaughterhouse works five days per week slaughtering a mean of 1700 pigs per day, obtaining more than 32000 carcasses per month. A total of 4000 carcasses were selected for this study according to their breed and sex in order to ensure a representative sample regarding fat thickness. Those carcasses were selected according to their sex: 60.6% females, 19.4% entire males and 20.0% castrated males and according to their genetics: 51.9% (Large White × Landrace) × Piétrain, 38.3% were (Large White x Landrace) x

Duroc and 9.8% **(**Large White x Landrace) x (Duroc x Landrace**).** Table 1.1 shows the mean weight of the cold carcass and the fat thickness of the pigs according to the breed and sex. Fat thickness parameter is given by the ultrasound AutoFom-III system and corresponds to the parameter F34 that is described as the fat thickness at 60 mm in the mid-line between the 3rd and the 4th last rib.

Table 1.1. The cold carcass weight (mean ± s.d; kg) and the fat thickness at 60 mm in the mid-line between the 3rd and the 4th last rib (mean ± s.d.; mm) of 4000 carcass according to breed and sex.

BREED	**n**	**WEIGHT (mean ± s.d; kg)**	**FAT THICKNESS (mean ± s.d.; mm)**
(Large White × Landrace) × Piétrain	2077	81.80 ± 8.16	15.39 ± 4.10
(Large White x Landrace) x Duroc	1531	93.76 ± 10.69	24.55 ± 5.56
(Large White x Landrace) x (Duroc x Landrace**)**	392	85.92 ± 9.02	18.60 ± 5.20
SEX			
Female	2289	85.49 ±10.00	17.59 ± 5.63
Castrated	1315	90.97 ± 11.54	23.51 ± 6.19
Entire male	396	80.38 ± 7.91	14.29 ± 3.22

Pigs were slaughtered after stunning with CO_2 (90%) for 2 min. After scalding they were totally monitored using the ultrasound AutoFom-III system. Then pigs were eviscerated and splitted according to standard commercial procedures using an automatic robotic system. After that, the two half-carcasses were weighted and an experimented operator visually determined the sex of the pig (female, entire male or castrated male) and classified the left half carcass according to minimal fat depth over muscle *gluteus medius* which is shown in Fig. 1.1. Classes were established based on the measures shown in Table 1.2. The operator had a pattern, based on these classes, that was used to visually compare and determine in which of the four ham classes (HC) each ham was classified.

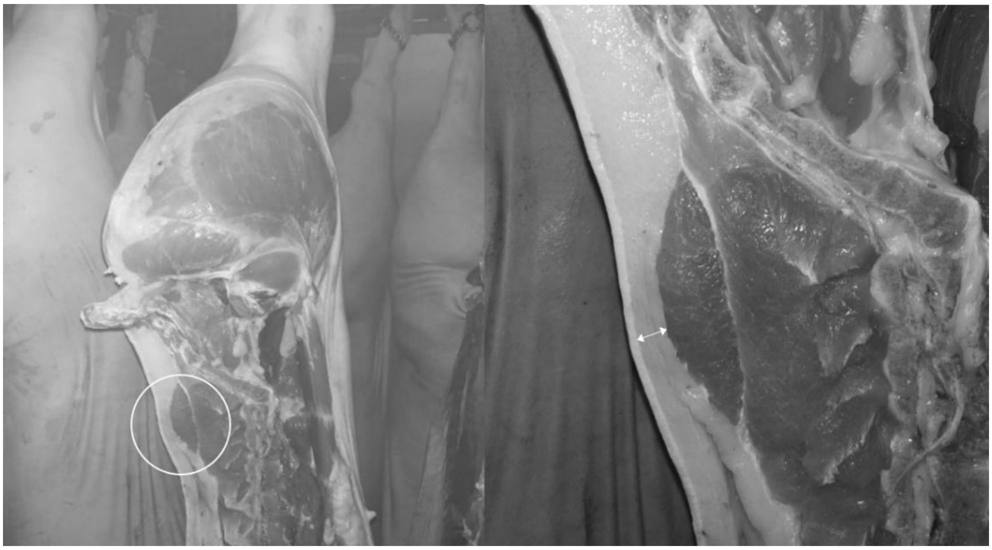

Fig. 1.1 Representation of the section used and the measure performed by an expert operator to measure the minimal fat thickness over muscle gluteus medius to obtain the classification target.

Table 1-2 Carcass classification according to minimal fat thickness over muscle gluteus medius based on a metrical measure with a ruler

Ham_Class (HC)	Fat depth (mm)
(1)- Thin	<10
(2)- Standard	Between < 10 and 15
(3)- Semi-fat	Between < 15 and 20
(4)- Fat	> 20

1.2.2. Dataset predictors

AutoFom-III predicts carcass LMP and seven other variables (Table 1.3) from 48 parameters obtained from the scanning. Nevertheless, a more accurate handmade classification process of the ham is required for commercial purposes. With the aim of improving classification rates the eight estimations provided by AutoFom-III, that are going to be used as predictors, are complemented with three more predictors obtained in the production line (sex, breed, and weight) (Table 1.3). The extended set of 11 predictors was used as the input of automatic classification systems applying pattern recognition techniques to assess different classifiers.

Table 1-3 The eleven predictors used as the input of automatic classification systems.

Predictor	Description
	Autofom III
LMP	Lean Meat Percentage
F34	According to the official formula, the subcutaneous fat thickness at 60 m in the mid-line between the 3^{rd} and the 4^{th} last rib. (mm)
M34	According to the official formula, muscle thickness at 60 mm in the mid-l from the 3^{rd} to the 4^{th} last rib. (mm)
F_GM1	The minimum subcutaneous fat plus skin thickness measured with a rul over the muscle *Gluteus medius* (mm)
F_GM2	The thickness of the subcutaneous fat plus skin measured with a ruler, perpendicularly to the skin, at the cranial part of muscle *Gluteus medius* (mm)
WGT_H	Total weight of the ham (kg)
WGT_HWB	Ham's weight without bone (kg)
WGT_HLM	Total weight of the lean meat of the ham (kg)
	Production line
SEX	Sex of animals (females, entire males and castrated males)
BREED	Crossbreed ((Large White x Landrace) x Pietrain, , (Large White x Landrace) x Duroc, and (Large White x Landrace) x (Duroc x Landrace))
WGT	Cold carcass weight (kg)

Finally, the HC parameter (1, 2, 3 or 4; see Table 1.2) used as a response was scored by an expert operator and is referred to the manual metrical measure to classify hams according to the thickness of the fat at the point shown in Fig.1.1.

1.2.3. Predictors and classifiers evaluated

A preliminary study was performed to evaluate the potential of each predictor individually to forecast the HC classification. Therefore, each single predictor was only considered to feed each of the classifiers to obtain the response. All classifiers were evaluated in terms of the accuracy which is defined as the number of correct predictions divided by the number of total predictions.

Moreover, the impact in the prediction of HC when taking different combinations of predictors as inputs in the classifier was also assessed in terms of the accuracy. The aim of this assessment was to compare the predictability of the classifiers when trained with only the single input LMP, and when other predictors are incorporated, such as the combinations of LMP and SEX or LMP and BREED (see Table 1.4) for all the combinations. These combinations were chosen according to the different slaughterhouse scenarios described below.

Table 1-4. Predictors included in each data set.

	Predictors used as inputs										
Datasets	LMP[1]	SEX[2]	WGT[3]	BREED[4]	F34[5]	M34[6]	F_GM1[7]	F_GM2[8]	WGT_H[9]	WGT_HWB	WGT_ HLM
D1	X										
D2	X	X									
D3	X		X								
D4	X			X							
D5	X	X	X	X							
D6	X	X	X	X	X	X					
D7	X	X	X	X	X	X	X	X	X	X	X

[1]LMP (Lean Meat Percentage); [2]SEX (females, entire males and castrated males); [3]WGT (warm carcass weight); [4]BREED ((Large White x Landrace) x Pietrain, (Large White x Landrace) x Duroc and (Large White x Landrace) x (Duroc x Landrace)); [5]F34 (subcutaneous fat thickness at 60 mm in the mid-line between the 3[rd] and the 4[th] last rib); [6]M34 (loin depth in mmmeasured at 60 mm from the midline between the 3[rd] and the 4[th] last rib); [7]F_GM1 (minimum subcutaneous fat plus skin thickness measured with a ruler over the muscle *Gluteus medius*); [8]F_GM2 (thickness of the subcutaneous fat plus skin measured with a ruler, perpendicularly to the skin, at the cranial part of muscle *Gluteus medius*); [9]WGT_H (total weight of the ham); [10]WGT_HWB (ham's weight without bone); [11]WGT_HLM(total weight of the lean meat of the ham).

According to the Commission Delegated Regulation (EU) 2017/1182, it is mandatory in all the slaughterhouses to classify pig carcasses by means of its LMP. Therefore the combination D1 (Table 1.4) is available in the production line of all slaughterhouses.

As more procedures are added in the slaughtering line, more predictors could be obtained in real-time such as SEX, BREED, and

WGT. Those additional predictors can be incorporated as inputs in the classifiers, as it has been done from D2 to D5.

Combination D6 considers the addition of predictors F34 and M34 that are provided by AutoFOM III. These predictors have been chosen because they can be assessed using other classification systems like Fat-O-Meat'er- FOM (Kempster, Chadwick, & Jones, 1985). Finally, D7 takes all additional information given by AutoFOM III (predictors F_GM1, F_GM2, WGT_H, WGT_HWB, and WGT_HTL) (Table 1.4.).

1.2.4. Statistical analysis

To train each classifier four sets of 1000 samples of each HC class were randomly selected from the total of 31188 ones to form a balanced group of 4000 samples. Afterwards, to prevent the classifier overfitting, a *5-Fold cross-validation* method was used (Bishop, 2006) dividing the dataset into *5* subsets, and for *5* times one of the *5* subsets was used as test set and the other *4* subsets get together to form a training set and the average error across all *5* trials was computed.

All classifiers were evaluated in terms of the accuracy (number of correct predictions divided by the number of total predictions).

A set of well-known classifier techniques was evaluated (Bishop, 2006): (1) Decision Trees: this type of algorithm is based on the construction of an automatic diagram of branches that appear according to the available data and the specific weight of each parameter. This algorithm was used with 4, 20 and 100 maximum split-levels; (2) Support Vector Machines (SVM): a discriminative classifier that separates classes by a hyperplane. The SVM algorithm is based on finding the optimal separating hyperplane that gives the largest minimum distance between the classes of the training data. This algorithm was used with four different kernels - linear, quadratic, cubic and Gaussian (Burges, 1998; Vapnik &

Chervonenkis, 1964); (3) K-Nearest Neighbour Classifiers (*K*-NN): a non-parametric supervised classifier based on the comparison of a sample against the *K* samples which most resemble assigning the most abundant class (Cover & Hart, 1967). This algorithm was used with six different configurations; (4) Discriminant Analysis with linear (Balakrishnama & Ganapathiraju, 1998; Fisher, 1936) and quadratic configurations based on finding a linear or quadratic combination of parameters that characterise or separates two or more classes.

MATLAB and Signal Processing Toolbox™ (Matlab R2016b; The MathWorks, Inc, 1988–2016) have been used to develop and test all the models and algorithms.

1.3. Results and discussion

Table 1.5 shows the accuracy of 17 classification models when a single predictor is taken as input. These classification models allow interpreting the results as a measure of the impact that each predictor by itself has in the forecast. Accuracy oscillates between 15 and 68% depending on the predictor and the type of classifier. Predictors F_GM1 and F_GM2 obtain the best results of accuracy in most of the classifiers, outperforming the results obtained by LMP. F34 also achieves good results regarding accuracy, however, in this case, the results are more dependent on the classifier type. Those results were foreseeable as predictors F_GM1, F_GM2 and F34 provide information about a direct measure of fat thickness in two points of the ham and in one point of the loin, respectivily. Indeed, they are physically related to the handmade measure taken by an expert operator who assigns the HC class. On the other hand, predictors such as SEX, WGT and BREED can be good predictors to classify the hams correctly but largely depends on the type of classifier.

The highest and the lowest accuracy values for each predictors' dataset are presented in bold and underlined, respectively. The best results of predictors F_GM1, F34, F_GM2 and LMP predicted the HC class with an accuracy between 63 and 68%. Moreover, predictors BREED, WGT and SEX predicted the HC class with an accuracy between 42 and 48%. Finally, the rest of predictors, had an accuracy below 37%.

In general, SVM Medium Gaussian or Coarse Gaussian or the Fine worked better when predictors are lean or fat parameters while SVM Cubic is one of the worst. This result persists in all predictors used but the interpretation about the relation of SVM kernels and the dataset is not clear.

When weight predictors are used, linear or quadratic discriminant analysis, and also Medium Gaussian, Coarse Gaussian and fine SVM produce the highest accuracy. These results suggest that continuous variables, such as the weight, improve the accuracy of more complex algorithms while categorical variables fits better with more simple algorithms. Sex and breed have higher accuracy when decision trees and SVM are used and discriminant analysis for breed. We can hypothesize than sex and breed obtain higher accuracy in decisions trees because, in the dataset, they are only three breed classes (Table 1.1). According to the results of (Gispert et al 2007), there is a clear relation between breed and SFT that could be easily formalized in simple decision trees. Similar relations have been found for sex (Font-i-Furnols et al., 2012; Gispert et al., 2010). The lowest accuracy is for the kNN approach. We can observe that for the classification of ham is usually more relevant breed than weight, and in turn, weight than sex.

Table 1-5. The Accuracy (in percentage) to predict the Ham Classification (HC) based on the thickness of the subcutaneous fat of the ham for each classifier when a single predictor is considered.

	Predictors										
Classifiers	**LMP[1]**	**F34[5]**	**M34[6]**	**F_GM1[7]**	**F_GM2[8]**	**WGT_H[9]**	**WGT HWB[10]**	**WGT HTL[11]**	**SEX[2]**	**WGT[3]**	**BREED[4]**
Decision Trees											
Simple	62	**65**	**36**	**68**	63	33	33	**32**	**42**	**44**	**48**
Medium	61	**65**	**36**	67	64	33	32	**32**	**42**	43	**48**
Complex	61	63	**36**	65	62	32	32	29	**42**	42	**48**
Support Vector Machines											
Linear	52	58	27	59	51	28	27	26	40	35	47
Quadratic	31	38	*25*	40	44	25	25	26	**42**	27	**48**
Cubic	15	19	25	35	22	24	23	24	**42**	19	**48**
Fine	**63**	**65**	**36**	**68**	**65**	34	33	**32**	**42**	**44**	**48**
Medium Gaussian	**63**	**65**	**36**	**68**	**65**	35	33	**32**	**42**	**44**	**48**
Coarse Gaussian	**63**	**65**	**36**	**68**	64	34	33	**32**	**42**	**44**	**48**
K-Nearest Neighbours											
Fine	36	53	30	57	53	27	28	25	25	26	25
Medium	58	62	33	65	63	31	32	28	25	33	25
Coarse	62	**65**	35	**68**	64	32	32	31	27	41	25
Cosine	25	25	25	25	25	25	25	25	25	25	25
Cubic	58	61	33	65	63	32	32	28	25	32	25
Weighted	57	56	31	60	55	29	30	27	*25*	32	*25*
Discriminant analysis											
Linear	62	**65**	**36**	**68**	64	34	33	**32**	34	**44**	**48**
Quadratic	61	64	**36**	**68**	63	**35**	**34**	**32**	39	**44**	**48**

[1]LMP (Lean Meat Percentage); [2]SEX (Females, entire males and castrated males); [3]WGT (warm carcass weight); [4]BREED ((Large White x Landrace) x Pietrain, (Large White x Landrace) x Duroc and (Large White x Landrace) x (Duroc x Landrace)); [5]F34 (subcutaneous fat thikness at 60 mm in the mid-line between the 3rd and the 4th last rib); [6]M34 (loin depth in mm measured at 60 mm from the midline between the 3rd and the 4th last rib); [7]F_GM1 (minimum subcutaneous fat plus skin thickness measured with a ruler over the muscle *Gluteus medius*); [8]F_GM2 (thickness of the subcutaneous fat plus skin measured with a ruler, perpendicularly to the skin, at the cranial part of muscle *Gluteus medius*); [9]WGT_H (total weight of the ham); [10]WGT_HWB (ham's weight without bone); [11]WGT_HLM(total weight of the lean meat of the ham). In bold he highest value for each dataset; Underlined lowest value for each dataset.

Table 1.6 shows the accuracy of each classifier according to the data set configurations that are more commonly available in different slaughterhouse scenarios, as described in section 2.3, Table 1.4. As commented in section 2.4 classifiers were obtained and validated with cross validation with the 4000 carcasses. In addition, although the 27188 were a non-balanced data set in terms of HC, (i.e. 16920 (thin), 6074 (standard), 4003 (semi-fat) and 191 (fat)) the classifiers were also validated using this dataset and accuracy of the results was similar to the obtained by cross validation (data not shown).

Table 1-6. Accuracy (in percentage) of each classification model with different dataset configurations1 used to train models.

	Datasets						
Classifiers	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**
Decision Trees							
Simple tree	62	62	63	62	64	65	68
Medium tree	61	64	65	62	67	68	70
Complex tree	61	63	64	61	68	67	68
Support Vector Machines							
Linear	52	61	**67**	**63**	**69**	**71**	71
Quadratic	31	45	60	49	68	**71**	72
Cubic	<u>15</u>	<u>32</u>	<u>33</u>	<u>37</u>	68	**71**	69
Fine	**63**	**65**	66	**63**	68	69	69
Medium Gaussian	**63**	**65**	**67**	**63**	**69**	70	**73**
Coarse Gaussian	**63**	64	**67**	**63**	68	**71**	71
K-Nearest Neighbours							
Fine	36	43	56	42	59	<u>61</u>	<u>62</u>
Medium	58	61	64	58	65	66	68
Coarse	62	64	56	62	65	68	67
Cosine	25	61	56	58	65	68	68
Cubic	58	61	65	57	65	68	68
Weighted	57	59	60	**63**	63	67	68
Discriminant Analysis							
Linear	62	56	**67**	59	64	67	70
Quadratic	62	58	65	54	<u>55</u>	63	66

[I]In bold the highest value for each dataset; Underlined the lowest value for each dataset.
[1] See Table 2 for description of the inputs included as predictors in each dataset studied (from D1 to D7).

The first column shows the results obtained using LMP as a single predictor. The highest value (stood out in bold, Table 1.6) of the different classifiers for dataset configurations. D2 and D3 show a positive impact on most of the classifiers accuracy due to the incorporation of SEX and WGT predictors, respectively, compared with D1. Moreover, dataset configuration D4, in which BREED predictor has been incorporated, the accuracy improves just in some of the classifiers, such as SVM Linear and KNN Cosine. Predictor WGT seems to better complement LMP than SEX and BREED according to results obtained by Latorre, García-Belenguer, & Ariño (2008).

As a general rule, SVM Coarse, SVM Medium Gaussian and SVM Fine obtain the highest accuracy when only one or two predictors are used (D1 to D4) compared with the other classifiers. Moreover, when more predictors are used, all the SVM classifiers produce better results than the other classifier techniques. In addition, the more predictors are added, the better results are obtained with the most sophisticated classifiers, such as SVMs with complex kernels.

When SEX, WGT and BREED predictors complement LMP (D5) the accuracy of SVM Medium Gaussian, one of the classifiers with the highest accuracy in D1, increases a 6%, obtaining an accuracy value of 69%. Furthermore, the SVM Linear with D5, also obtain an accuracy value of 69% increasing by 17% with respect to D1.

D6 dataset configuration incorporates to D5 predictors F34 and M34 obtained by Autofom. Configuration D7 has all available predictors (see section 2.2), obtained through the use of Autofom and intrinsic characteristics of the animal. In configurations D6 and D7, the classifiers obtain a percentage of accuracy between 61 and 73%. As expected, D7 configuration obtains the best performance. Regarding the classifiers, the SVM Medium Gaussian reached the best result with a percentage of accuracy of 73%.

When comparing models obtained from datasets D6 and D7, in average, there is a 1.0% of prediction improvement. It is suggested that the improvement is not greater because the added parameters are closely correlated with the previous ones. For instance, the five new predictors (F_GM1, FGM2, WGT_H, WGT_HWB, WGT_HLM) introduced in the models with input dataset D7 are highly correlated with predictors WGT and/or F34, present in dataset D6. However, although an increase of 1.0% does not represent a great improvement in terms of percentage of success, it can mean a significantly improvement in the benefits of a company. Misclassifications of a ham in a lower category, in terms of subcutaneous fat, could incur in losses of more than 30% in the final sale price.

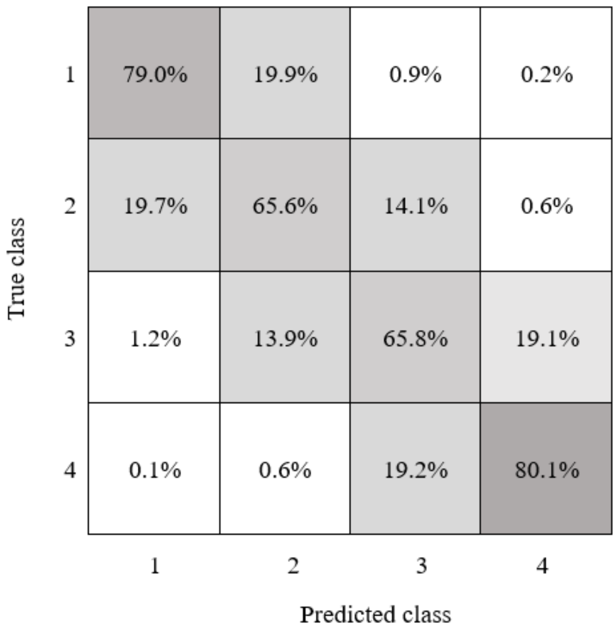

Fig. 1.2 Confusion Matrix of the best accuracy models obtained using SVM medium Gaussian model trained with all data available (D7). The results are given in percentatge.

Fig.1.2 shows the confusion matrix obtained by SVM Medium Gaussian model developed using D7. The accuracy of HC classes 1 (79.0%) and 4 (80.1%) are higher than the accuracy of HC classes 2

(65.6%) and 3 (65.8%). When classes based on a metric threshold are used, extreme classes tend to be better classified.

The percentage of samples that are incorrectly classified into one of the adjacent categories varies between 13.9%-19.9% (Fig. 1.2.). It should be noted that some of these samples fall very close to the decision thresholds and, in those cases, the classification is particularly difficult.

Moreover, only less than 3.6% of the samples are misclassified in not adjacent categories. Indeed, it can be concluded that 96.7% of the 27.4% of misclassified samples correspond to samples classified into adjacent categories.

As explained before, all the models are developed in order to predict the classification of the hams by an expert operator. Indeed, in this study the human classification methodology is used as "golden standard" despite the fact that this methodology presents some difficulties such as operator fatigue (Font-i-Furnols et al., 2016; Olsen et al., 2007) but also the evaluation of the fat thickness after the carcass being split down by an industrial robot (the carcasses are not precisely split down in the same way). Therefore, misclassifications of the models do not always mean that the model is classifying wrong, they are just explaining that the model classification does not match with the human classification.

Nowadays, the SVM Medium Gaussian model is applied in MAFRICA S.A. slaughterhouse. It is observed an accuracy improvement which is not currently quantified. Our working hypothesis is that automatic classification improves manual classification because decision making is objective and operator fatigue are eliminated.

1.4. Conclusions

Pattern recognition models, based on data usually available on slaughterhouses, can be used to classify the hams according to the thickness of the subcutaneous fat, and this classification can emulate the manual system with an effectivity of 73%. This result suggests that pattern recognition models can be a useful online tool to increase slaughterhouses' benefits because more accurate classification increases optimization of the ham processing.

Acknowledgements

"Financial support from the Agency for Management of University and Research Grants (AGAUR) of the Catalan Government to Gerard Masferrer is grateful". CERCA Programme (Generalitat de Catalunya) is also acknowledged.

References

Balakrishnama, S., & Ganapathiraju, A. (1998). Linear Discriminant Analysis a Brief Tutorial. Compute, 11, 1–9. http://www.isip.piconepress.com/publications/reports/1998/isip/lda/

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Pattern Recognition (Vol. 4). http://doi.org/10.1117/1.2819119

Bosi, P., Russo, V., & Paolo, B. (2004). The production of the heavy pig for high quality processed products. J.ANIM.SCI, 3(21), 309–32145. http://doi.org/10.4081/ijas.2004.309

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2(2), 121–167. http://doi.org/10.1023/A:1009715923555

Candek-Potokar, M., & Skrlep, M. (2012). Factors in pig production that impact the quality of dry-cured ham: a review. Animal : An International Journal of Animal Bioscience, 6(2), 327–38. http://doi.org/10.1017/S1751731111001625

Carabús, A., Gispert, M., Brun, A., Rodríguez, P., & Font-i-Furnols, M. (2014). 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. http://doi.org/10.1016/j.livsci.2014.10.005

Cilla, I., Altarriba, J., Guerrero, L., Gispert, M., Martínez, L., Moreno, C., … Roncalés, P. (2006). Effect of different Duroc line sires on carcass composition, meat quality and dry-cured ham acceptability. Meat Science, 72(2), 252–260. http://doi.org/10.1016/j.meatsci.2005.07.010

Cisneros, F., Ellis, M., & McKeith, F. (1996). Influence of slaughter weight on growth and carcass characteristics, commercial cutting and curing yields, and meat quality of barrows and gilts from two genotypes. Journal of Animal. Retrieved from https://dl.sciencesocieties.org/publications/jas/abstracts/74/5/925

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 4.7.17, L171/74-L171-99.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. http://doi.org/10.1109/TIT.1967.1053964

Daumas, G. (2011). Non-electronic techniques to classify pig carcasses in small slaughterhouses. In Second International Virtual Conference on pork Quality. Novembere, 05to Decembere (p. p 06-2001).

Dirinck, P., Van Opstaele, F., & Vandendriessche, F. (1997). Flavour differences between northern and southern European cured hams. Food Chemistry, 59(4), 511–521. http://doi.org/10.1016/S0308-8146(97)00012-5

Edwards, D. B., Bates, R. O., & Osburn, W. N. (2003). Evaluation of Duroc- vs. Pietrain-sired pigs for carcass and meat quality measures. Journal of Animal Science, 81(8), 1895. http://doi.org/10.2527/2003.8181895x

Fàbrega, E., Gispert, M., Tibau, J., Hortós, M., Oliver, M. A., & Font i Furnols, M. (2011). Effect of housing system, slaughter weight and slaughter strategy on carcass and meat quality, sex organ development and androstenone and skatole levels in Duroc finished entire male pigs. Meat Science, 89(4), 434–439. http://doi.org/10.1016/j.meatsci.2011.05.009

Fisher, R. A. (1936). The use of multiple measurements in taxanomic problems. Annals of Eugenics, 7(2), 179–188. http://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Font-i-Furnols, M., Čandek-Potokar, M., Daumas, G., Gispert, M., Judas, M., & Seynaeve, M. (2016). Comparison of national ZP equations for lean meat percentage assessment in SEUROP pig classification. Meat Science, 113, 1–8. http://doi.org/10.1016/j.meatsci.2015.11.004

Font-i-Furnols, M., Gispert, M., Soler, J., Diaz, M., Garcia-Regueiro, J. A., Diaz, I., & Pearce, M. C. (2012). Effect of vaccination against gonadotrophin-releasing factor on growth performance, carcass, meat and fat quality of male Duroc pigs for dry-cured ham production. Meat Science, 91(2), 148–154. http://doi.org/10.1016/j.meatsci.2012.01.008

Font i Furnols, M., & Gispert, M. (2009). Comparison of different devices for predicting the lean meat percentage of pig carcasses. Meat Science, 83(3), 443–446. http://doi.org/10.1016/j.meatsci.2009.06.018

Gispert, M., Àngels Oliver, M., Velarde, A., Suarez, P., Pérez, J., & Font i Furnols, M. (2010). Carcass and meat quality characteristics of immunocastrated male, surgically castrated male, entire male and female pigs. Meat Science, 85(4), 664–670.http://doi.org/10.1016/j.meatsci.2010.03.021

Gispert, M., Font i Furnols, M., Gil, M., Velarde, A., Diestre, A., Carrión, D., … Plastow, G. S. (2007). Relationships between carcass quality parameters and genetic types. Meat Science, 77(3), 397–404. http://doi.org/10.1016/j.meatsci.2007.04.006

Kempster, A. J., Chadwick, J. P., & Jones, D. W. (1985). An evaluation of the Hennessy grading probe and the SFK Fat-O-Meater for use in pig carcass classification and grading. Animal Production, 40(2), 323–329. http://doi.org/10.1017/S0003356100025447

Latorre, M. A., García-Belenguer, E., & Ariño, L. (2008). The effects of sex and slaughter weight on growth performance and carcass traits of pigs intended for dry-cured ham from Teruel (Spain). Journal of Animal Science, 86(8), 1933–1942. http://doi.org/10.2527/jas.2007-0764

Lucas, D., Brun, A., Gispert, M., Carabús, A., Soler, J., Tibau, J., & Font-i-Furnols, M. (2017). Relationship between pig carcass characteristics measured in live pigs or carcasses with Piglog, Fat-o-Meat'er and computed tomography. Livestock Science, 197, 88–95. http://doi.org/10.1016/j.livsci.2017.01.010

Nanni Costa, L., Stelletta, C., Cannizzo, C., Gianesella, M., Lo Fiego, D. P., & Morgante, M. (2010). The use of thermography on the slaughter-line for the assessment of pork and raw ham quality. Italian Journal of Animal Science, 6(1s). http://doi.org/10.4081/ijas.2007.1s.704

Olsen, E. V., Candek-Potokar, M., Oksama, M., Kien, S., Lisiak, D., & Busk, H. (2007). On-line measurements in pig carcass classification: Repeatability and variation caused by the operator and the copy of instrument. Meat Science, 75(1), 29–38. http://doi.org/10.1016/j.meatsci.2006.06.011

Realini, C. E., Duran-Montgé, P., Lizardo, R., Gispert, M., Oliver, M. A., & Esteve-Garcia, E. (2010). Effect of source of dietary fat on pig performance, carcass characteristics and carcass fat content, distribution and fatty acid composition. Meat Science, 85(4), 606–612. http://doi.org/10.1016/j.meatsci.2010.03.011

Tous, N., Lizardo, R., Vilà, B., Gispert, M., Font-I-Furnols, M., & Esteve-Garcia, E. (2014). Effect of reducing dietary protein and lysine on growth performance, carcass characteristics, intramuscular fat, and fatty acid profile of finishing barrows 1. J. Anim. Sci, 92, 129–140. http://doi.org/10.2527/jas2012-6222

Vapnik, V. N., & Chervonenkis, A. (1964). A note on one class of perceptrons. Automation and Remote Control, 25(6), 937–945. Retrieved from https://scholar.google.es/scholar?q=Vladimir+Vapnik+Alexey+Ya.+Chervonenkis&hl=ca&as_sdt=0%2C5&as_ylo=1960&as_yhi=1965

Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Sheard, P. R., Richardson, R. I., … Whittington, F. M. (2008). Fat deposition, fatty acid composition and meat quality: A review. Meat Science. http://doi.org/10.1016/j.meatsci.2007.07.019

Wood, J. D., Enser, M., Whittington, F. M., Moncrieff, C. B., & Kempster, A. J. (1989). Wood et al., 1989 Backfat composition in pigs Differences between fat thickness groups and sexes..pdf. Livestock Production Science Elsevier Science Publishers B.V.

Wood, J. D., Nute, G. R., Richardson, R. I., Whittington, F. M., Southwood, O., Plastow, G., … Chang, K. C. (2004). Effects of breed, diet and muscle on fat deposition and eating quality in pigs. Meat Science, 67(4), 651–667. http://doi.org/10.1016/j.meatsci.2004.01.0

2. AUTOMATIC HAM CLASSIFICATION METHOD BASED ON SUPPORT VECTOR MACHINE MODEL INCREASES ACCURACY AND BENEFITS COMPARED TO MANUAL CLASSIFICATION

Automatic ham classification method based on support vector machine model increases accuracy and benefits compared to manual classification

Gerard Masferrer[a,b], Ricard Carreras[b], Maria Font-i-Furnols[c], Marina Gispert[c], Moises Serra[a], Pere Marti-Puig[a]

[a]*Information and Digital Technology Department, UVic-UCC, Vic (Barcelona)*

[b]*Mafrica.SA, Paratge Can Canals Nou, S/N 08250 Sant Joan de Vilatorrada (BCN)*

[c]*IRTA - Product Quality, 17121 Monells, Girona, Spain*

Abstract

The thickness of the subcutaneous fat (SFT) is a very important parameter in the ham, since determines the process the ham will be submitted. This study compares two methods to predict the SFT in slaughter line: an automatic system using an SVM model (Support Vector Machine) and a manual measurement of the fat carried out by an experienced operator, in terms of accuracy and economic benefit. These two methods were compared to the golden standard obtained by measuring SFT with a ruler in a sample of 400 hams equally distributed within each SFT class. The results show that the SFT prediction made by the SVM model achieves an accuracy of 75.3%, which represents an improvement of 5.5% compared to the manual measurement. Regarding economic benefits, SVM model can increase them between 12-17%. It can be concluded that the classification using SVM is more accurate than the one performed manually with an increase of the economic benefit for sorting.

Keywords dry-cured hams; ham-fat grading; subcutaneous fat thickness; pattern recognition; sorting

2.1. Introduction

The thickness of subcutaneous fat (SFT) is one of the most critical parameters in hams for several reasons. Indeed, the thickness of the fat usually determines the process to which the ham will be subjected: dry-curing, cooking or the processing of the raw meat (Bosi, Russo, & Paolo, 2004). Moreover, SFT is particularly significant in the dry-curing process, as it is one of the critical factors determining the final quality of the product (Candek-Potokar & Skrlep, 2012). The SFT and the weight mostly determine the amount of salt and other ingredients necessary for the dry-curing process (Škrlep et al., 2016) and the curing time (Buscailhon, Gandemer, & Monin, 1994; Marriott, Graham, & Claus, 1992; Toldrá & Flores, 1998; Toldrá, Flores, & Sanz, 1997). Besides, SFT measure is essential to determine the yield in the production of raw meat, which in turn determines the lean percentage of the piece.

It is possible to measure SFT once the green ham has been shaped, however it would be interesting to estimate SFT measure on-line in order to classify the ham before being processed (Masferrer et al. 2018). Optimization of the industrial processes in slaughterhouses is essential to become more competitive. A correct classification of the carcass and the ham on-line can lead to a substantial increase in the slaughterhouse yield. Indeed, it reduces the number of reprocessed primal cuts and allows the linearization of the processes in the cutting plant according to the characteristics of the batches, such as SFT, weight of the ham or breed.

For the online classification of the carcasses and, consequently, the hams, the carcass weight and the lean meat percentage (LMP) are usually used, as they are mandatory data that slaughterhouses must measure according to the Commission Delegated Regulation (EU) 2017/1182. Although, the results obtained using those parameters are positively correlated with ham SFT (Gispert et al., 2007; Pulkrábek, Pavlík, & Valis, 2006), there is significant room for improvement. On the other hand, there are slaughterhouses where hams are specifically classified according to their own characteristics, such as SFT or LMP of the ham. Furthermore, some slaughterhouses use online predictors obtained from automatic or semi-automatic devices (Font i Furnols & Gispert, 2009), such as AutoFom and Fat-o-Meater or manual classification of the ham measuring the SFT employing a pattern according to ZP (Zwei-Punkte Messverfahren) measure or similar (Font-i-Furnols et al., 2016).

Pattern recognition systems are widely used in many fields (Bishop, 2006; Jain, Duin, & Jianchang Mao, 2000). One of the most commonly used algorithms is the Support Vector Machine (SVM). The SVM algorithm is based on finding a hyperplane of separation between different categories. These type of algorithms allow making predictions of categories, in our case the categories based on the SFT of the ham. Those algorithms could be a useful tool in the meat industry, as the amount of data collected from the entire production chain, including the farm and the slaughterhouse, can be significant.

Another relevant factor is the speed of the classification process, as the automatic algorithms, in addition to replacing manual work, allow sorting at high rates compared with manual classification. The objective of this study was to compare a SVM classification algorithm with a manual classification system, commonly used in commercial slaughterhouses, in order to classify hams according to their SFT. The SVM algorithm employs a middle Gaussian core, the best model obtained in Masferrer et al. (2018), which is trained with intrinsic data of the pigs (sex, weight and breed) and data predicted by AutoFom-III (Frontmatec Smørum A/S, Herlev, Denmark) (Brøndum, Egebo, Agerskov, & Busk, 1998). Furthermore, it is also an objective of this work to evaluate the economic effect of the accuracy on ham classification method for the meat industry.

2.2. Material and Methods

2.2.1. Animals and facilities

A total of 400 hams were selected from pigs slaughtered in the 13th of March 2018 at a commercial slaughterhouse (MAFRICA S.A.) located in Sant Joan de Vilatorrada, Catalonia, Spain (see section 2.2). The animals selected included three different genetic lines: (Large White × Landrace) × Piétrain, (Large White x Landrace) x Duroc and (Large White x Landrace) x (Duroc x Landrace). Animals came from farms, all of them less than 200 km far from the slaughterhouse and pigs were transported using trucks in groups of between 80 and 220 animals. Once in the slaughterhouse pigs rested into lairage pens between 2 and 4 hours before being slaughtered.

Pigs were slaughtered after stunning with CO_2 (90%) for 2 min. After pig scalding process, pigs were totally scanned using the ultrasound AutoFom-III system. Then pigs were eviscerated and automatically split according to standard commercial procedures using a robot. After that, the two half-carcasses were weighted and an experienced operator visually determined the sex of the pig (female, entire male or castrated male) and classified the half carcass as described in Masferrer et al. (2018), in order to normalize the classification. The left half carcass was always used to avoid possible errors produced by the robot cut deviation. For the Manual Classification (HC_M), SFT was measured with a ruler according to minimal fat depth over muscle *gluteus medius*. The following SFT thresholds were used: class HC1: < 9 mm; class HC2: between 9-12 mm; class HC3: between 13-19 mm and class HC4: > 19 mm. The thresholds used were determined by commercial requirements of the slaughterhouse where this study was carried out. HC_M was performed by one experienced operator who usually does the classification in the line.

With the measures obtained with Autofom III and including information about sex, breed and warm carcass weight (see Table.2.1), the ham class predicted by SVM algorithm (HC_SVM) was obtained (Masferrer et al, 2018).

Table 2-1. The eleven predictors used as the input of automatic classification system (SVM)

Predictor	Description
	Autofom III
LMP	Lean Meat Percentage
F34	According to the official formula, the subcutaneous fat thickness at 60 m in the mid-line between the 3rd and the 4th last rib. (mm)
M34	According to the official formula, muscle thickness at 60 mm in the mid-line between the 3rd and the 4th last rib. (mm)
F_GM1	The minimum subcutaneous fat plus skin thickness measured with a rul over the muscle *Gluteus medius*. (mm)
F_GM2	The thickness of the subcutaneous fat plus skin measured with a ruler, perpendicularly to the skin, at the cranial part of muscle *Gluteus medius* (mm)
WGT_H	Total weight of the ham. (kg)
WGT_HWB	Ham's weight without bone. (kg)
WGT_HLM	Total weight of the lean meat of the ham. (kg)
	Production line
SEX	Sex of animals (females, entire males and castrated males)
BREED	Crossbreed ((Large White x Landrace) x Pietrain, , (Large White x Landrace) x Duroc, and (Large White x Landrace) x (Duroc x Landrace))
WGT	Warm carcass weight (kg)

2.2.2. Ham shaping

Once the carcasses were pre-trimmed in the cutting room, they were refrigerated for 24 hours in a chilling room. When carcasses reached approximately 4ºC and the ham was extracted and processed to give the final shape and classified according to customer specifications (weight and SFT). Final shape process consisted of removing the tail, rounding off the bottom of the ham and lifting the leg (Fig. 2.1).

Fig. 2.1 Ham after final shape process

After the final shape process, 400 hams, one from each carcass were selected according to the SFT measured at that moment. To obtain this parameter an operator employed a ruler to measure the minimal SFT of the ham located in the central part of the muscle *gluteus medius*, perpendicular to the skin (Golden Standard measure), as shown in Fig.2.2. Those 400 hams were equally distributed in four SFT classes of 100 samples: HC1: < 9 mm; class HC2: between 9-12 mm; class HC3: between 13-19 mm and class HC4: > 19 mm). Hams were randomly measured until 100 samples were obtained for each category. When a class reaches 100 samples, no more hams were selected for that class. The measures obtained using this methodology were necessary to create the Ham Classification used as Golden Standard, (HC_GS) in order to assess the accuracy of the prediction of HC_M and HC_SVM classifications.

Fig. 2.2 Representation of the section and the ruler used to measure the minimal SFT of the ham located in the muscle gluteus medius.

The characteristics of the pigs included in this work are presented in Table 2.2. It shows the mean and the standard deviation of the warm carcass weight (kg) and the fat thickness (mm) of the evaluated carcasses according to breed and sex. Fat thickness parameter is given by the ultrasound AutoFom-III system and it corresponds to the parameter F34, which is described as the fat thickness at 60 mm in the mid-line between the 3rd and the 4th last ribs.

Table 2-2 Warm carcass weight and fat thickness at 60 mm in the mid-line between the 3rd and the 4th last rib of 400 carcasses according to breed and sex.

BREED	n	WEIGHT (mean ± s.d; kg)	FAT THICKNESS (mean ± s.d; mm)
(Large White × Landrace) × Piétrain	218	89.95 ± 8.12	15.13 ± 4.38
(Large White x Landrace) x Duroc	139	93.19 ± 9.46	23.71 ± 5.11
(Large White x Landrace) x (Duroc x Landrace)	43	94.48 ± 8.68	22.11 ± 9.36
SEX			
Female	205	91.95 ±8.88	16.25 ± 4.10
Castrated males	150	92.32 ± 9.54	24.49 ± 6.49
Entire males	45	87.23 ± 3.31	11.99 ± 2.14

2.2.3. Statistical analysis

The objective of this statistical evaluation was to compare the classification obtained with the manual classification (HC_M) and the automatic classification performed by the SVM, respectively. SVM refers to Gaussian Medium algorithm (HC_SVM), the best model obtained in Masferrer et al. (2018), to classify hams according to SFT. In order to evaluate the prediction of the ham class, the SFT of the finished hams was specifically measured for this study, in order to obtain the Golden Standard Ham Class (HC_GS) as described in the previous section.

To compare the results obtained with HC_M and HC_SVM a Cohen's kappa coefficient (k-values) was used to compare the classification performed by HC_M and HC_SVM methods and to compare HC_M and HC_SVM classifications with HC_GS, respectively. The guidelines developed by Landis & Koch (1977) were used to interpret the k-values: poor agreement ($k < 0.00$), soft agreement ($k = 0.00$-0.20), fair agreement ($k = 0.21$-0.40), moderate agreement ($k = 0.41$-0.60), substantial agreement ($k = 0.61$-0.80) and almost perfect agreement ($k = 0.81$-1).

The means and variances of SFT measured by HC_GS according to the category assigned by the classification systems (HC_M and HC_SVM) were calculated. T-Test and Bartlett's Test were carried out to assess the mean and the homogeneity of variances, respectively. The significance level was established at $p<0.05$. Descriptive data is presented with means of mm and the standard error (mean±SE).

Moreover, the accuracy of the prediction of HC_M and HC_SVM according with the results of HC_GS was calculated. The accuracy of the prediction is defined as the number of correct predictions divided by the number of total predictions. As a result, two confusion matrices were constructed using a cross-table with the real values (HC_GS) obtained by the "golden standard", and those provided by the operator (HC_M) and by the algorithm (HC_SVM), respectively. The accuracy was also calculated by breed and sex.

In order to obtain additional information (i.e. assess if better results could be obtained using Golden Standard measures to train SVM, instead of using measures obtained with the manual measurement on-line), the SVM was trained with the measures obtained to create the HC_GS. The same predictors were used (see Table 2.1) but using HC_GS as an independent variable (response). Moreover, the same SVM Medium Gaussian was used, and the training and verification phase was a 5-fold Cross-Validation.

MATLAB, Statistics and Machine Learning Toolbox™ (Matlab R2018a; The MathWorks, Inc, 1988-2019) have been used to develop and test all the models and algorithms.

Moreover, the economic impact using those classification methods on the slaughterhouse was analysed and compared between the HC_M and HC_SVM. This slaughterhouse slaughters approximately 500,000 pigs/year (8,000-10,000 pigs/week). The distribution of ham classes according to sales data of hams of 2017 was 54% as HC1, 28% as HC2, 13% as HC3 and 5% as HC4. The economic data taken as reference correspond to the slaughterhouse where this study was carried out (MAFRICA S.A.). The economic profit according to increase in price due ham classification has been taken into account. This increase in prices was estimated reflecting an optimistic and pessimistic increased price according to slaughterhouse commercial data.

Hams in category HC1 have no increase in price (i.e. +0 €/kg), hams in category HC2 have an increase of price between +0.03 and +0.10 €/kg, hams in category HC3 have an increase of price between +0.12 and +0.20 €/kg and hams in category HC4 have an increase of price between +0.15 and 0.30 €/kg. To estimate the weight of the hams, the average of the warm weight of the carcasses of the population of 2017 (i.e. 87 kg) was used, taking into account that the ham represents between 28-30% of the carcass (Cisneros, Ellis, McKeith, McCaw, & Fernando, 1996; Gispert et al., 2007).

An increase of economic value was only applied when the ham was correctly classified by the predictors. Moreover, when a ham was misclassified, it was considered that the cost of reprocess or reclassify had a similar cost to the increase of price that could be assigned.

2.3. Results and discussion

2.3.1. Comparisons of classification methods

The confusion matrices plot as a cross-table are shown in Fig. 2.3, where the rows correspond to the predicted class (HC_M and HC_SVM, respectively) and the columns correspond to the HC_GS class, that is, the Golden Standard. The diagonal cells correspond to samples that are correctly classified. The off-diagonal cells correspond to incorrectly classified samples. Both the number of samples and the percentage of the total number of samples are shown in each cell. The column on the right of the matrix shows the percentages of all the samples predicted to belong to each class that are correctly (positive predictive value) and incorrectly (false discovery rate) classified. The row at the bottom shows the percentages of all the samples belonging to each class that are correctly (true positive rate) and incorrectly (false negative rate) classified. The cell in the bottom right of the matrix shows the overall accuracy.

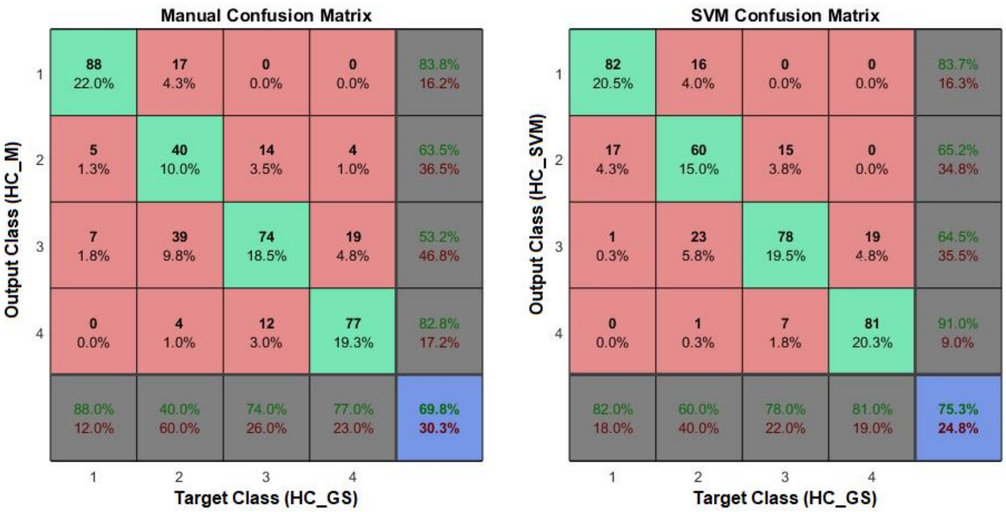

Fig. 2.3 Confusion matrices of Manual Ham Classification (HC_M) and Support Vector Machine Ham Classification (medium Gaussian kernel) (HC_SVM) compared with Golden Standard Ham classification (HC_GS). The results are given in number of observations and in perce of accuracy.

The accuracy of the prediction of HC_SVM was better than the HC_M (75.3% and 69.8%, respectively). Indeed, HC_SVM obtained better results in all classes except for HC1, where the HC_M obtained an 88% of correct predictions compared to the HC_SVM that obtained an 82%. This exception could be related to some Autofom III parameter used in the SVM algorithm. Specifically, the parameter F_GM1 (The minimum subcutaneous fat thickness over the muscle *Gluteus medius* (mm)) seems to be difficult to measure by AutoFom when SFT is very low. Indeed, Fig. 2.4 shows that hams with SFT lower than 6 mm according to HC_GS are overestimated by the F_GM1 parameter, (only eight hams have values below 6 mm of subcutaneous fat from 32 obtained in HC_GS). Moreover, Fig. 2.4 shows that F_GM1 estimates the central values of SFT with more precision than extreme values, especially underestimate higher SFT values.

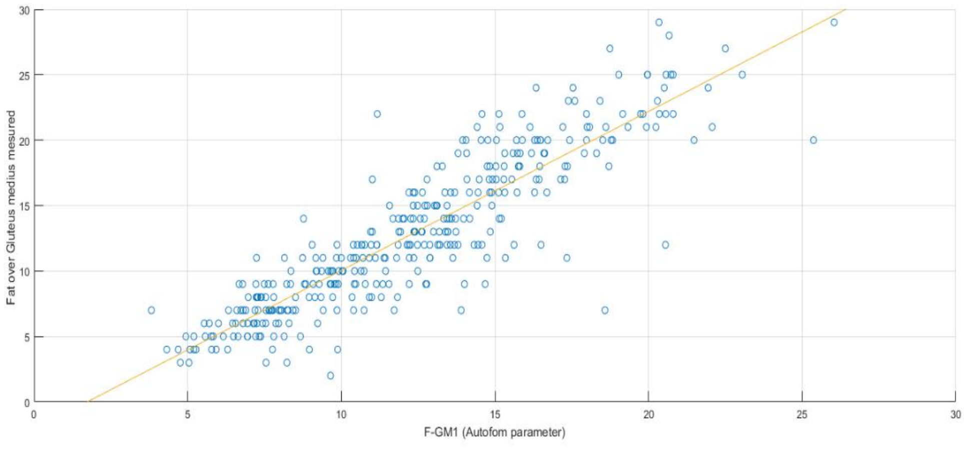

Fig. 2.4 Correlation between Golden Standard measure (mm of subcutaneous fat thickness of the ham after final shape process) and F-GM1 Autofom III parameter (mm of the minimum subcutaneous fat thickness over the muscle Gluteus medius).

Concerning the remaining HC2, HC3, and HC4 classes HC_SVM provides better predictions than HC_M improving the accuracy of prediction of HC3 and HC4 categories by a 4% and surprisingly by 20% in HC2. These results suggest that the operator tended to overestimate the class HC2, with a 9.8% of hams classified in HC3.

This tendency is also shown in Fig. 2.5 where the mean of the SFT estimated by HC_MC measures is higher than the mean of the SFT estimated by HC_SVM measures (12.34 ± 0.65 HC_M2 and 10.58 ± 0.25 in HC_SVM2; t-test p=0.007*)*. Moreover, it seems that the HC_M classification presented higher standard error within their group compared to HC_SVM classification (*p*= 0.000) (see Fig. 2.5). Indeed, the HC_M tend to overestimate (16.9%) in more cases than underestimate (13.6%) while the HC_SVM tend to overestimate as well as underestimate (12.5 and 12.6%, respectively). These results suggest that measurement methods could explain those differences. A possible explanation is related to fat state, when the operator measured SFT the carcass was hot and was not compacted as it was subjected vertically. Maybe those differences could be the reason for observing more deviation in HC_M measures (see Fig.2.5). While SFT was measured by Autofom the carcass, even hot, was compacted because this measure was recorded with the carcass with an horizontal position supported on a surface. Moreover, when the Golden Standard measure was taken the carcass had been cooling for 24 hours and the fat was compacted because of low temperatures. Differences between HC_GS and the two other classification methodologies, HC_M and HC_SVM, could also be related to fatty acid profile. According to St. John et al. (1987) and Warnants, Van Oeckel, & Boucqué (1996), as more content of unsaturated fatty acids more decreases fat firmness and is softer than fat with more content of saturated fatty acids which may affect the thickness of the fat. Although less fatty hams tend to be more unsaturated (Ruiz-Carrascal, Ventanas, Cava, Andrés, & García, 2000), contrary to what would be expected, fewer differences were found between HC_M and HC_SVM regarding less fat categories, i.e. HC1 and HC2. This result suggests that the effect on the softness of the fat, due to saturated fatty acid content, is less appreciable in less fat categories because of the lower content of fat.

Comparing the results showed on the right column of both matrices suggest that predictions of hams classified in categories HC3 and HC4 obtain better results in HC_SVM than in HC_M (64.5% vs. 53.2% in HC3 and 91.0% vs. 82.8% in HC4, respectively). While the results of the predictions of HC1 and HC2 are similar between HC_M and HC_SVM. These results have an impact on the economic benefits, since the categories HC3 and HC4 have a higher economic value than the other categories. These hams are intended for dry-curing processes and it is especially important to increase the percentage of true positives.

Comparing the incorrectly predicted samples between HC_M and HC_SVM is interesting to observe the dispersion of those incorrected samples. While the HC_SVM had the 0.6% of the incorrect samples not distributed to neighbouring classes, this result increased to 3.8% in the case of the HC_M. Indeed, this dispersion is also showed in Fig.2.5 where the standard error is presented for each class in both ham classifications (HC_M and HC_SVM). These cases can lead to unexpected production line problems that force the ham to be reprocessed offline, or being processed inefficiently. Usually, a certain percentage of hams are expected to be destined to neighbouring categories, but they are processed in a similar way, for example, HC3 or HC4 hams are usually destined for dry-curing processes while HC1 or HC2 hams are usually deboned to produce raw meat.

In addition to validate the results presented according to accuracy and confusion matrix a Cohen's kappa was used to take into account the possibility that the classification is produced by chance. Comparing the classification carried out by HC_M and HC_SVM, a moderate almost substantial agreement was found (k= 0.596). Although it is a good result, the agreement between the two classifications methods could be expected to be slightly higher, as the HC_SVM is a method created to emulate HC_M. Perhaps this result could be explained due to the number of samples used. In this study only 400 hams were analysed due to technical reasons (specifically to obtain the Golden Standard measures). On the other hand, the SVM algorithm used in the present study (HC_SVM) was trained with more than 30000 obtained samples, being more robust.

Moreover, when comparing both classification methods with the HC_GS, a substantial agreement (k = 0.670) was found for the HC_SVM method and a moderate almost substantial agreement (k = 0.597) was obtained for the HC_M. According to these results, it seems that the correct classifications with HC_SVM were slightly higher than with HC_M. Moreover, Fig.2.5 shows the mean and standard error in mm of SFT of the hams over the muscle *gluteus medius* calculated by the Golden Standard method for each class (HC1, HC2, HC3 and HC4) and for both methods (HC_M and HC_SVM). Indeed, this figure shows that the standard error of the measures obtained by HC_SVM was lower than the standard error of the values obtained by HC_M. And the results of Bartlett's test show a significant difference on variance between methods in HC2 and HC3 (p= 0.000 and p= 0.038, respectively).

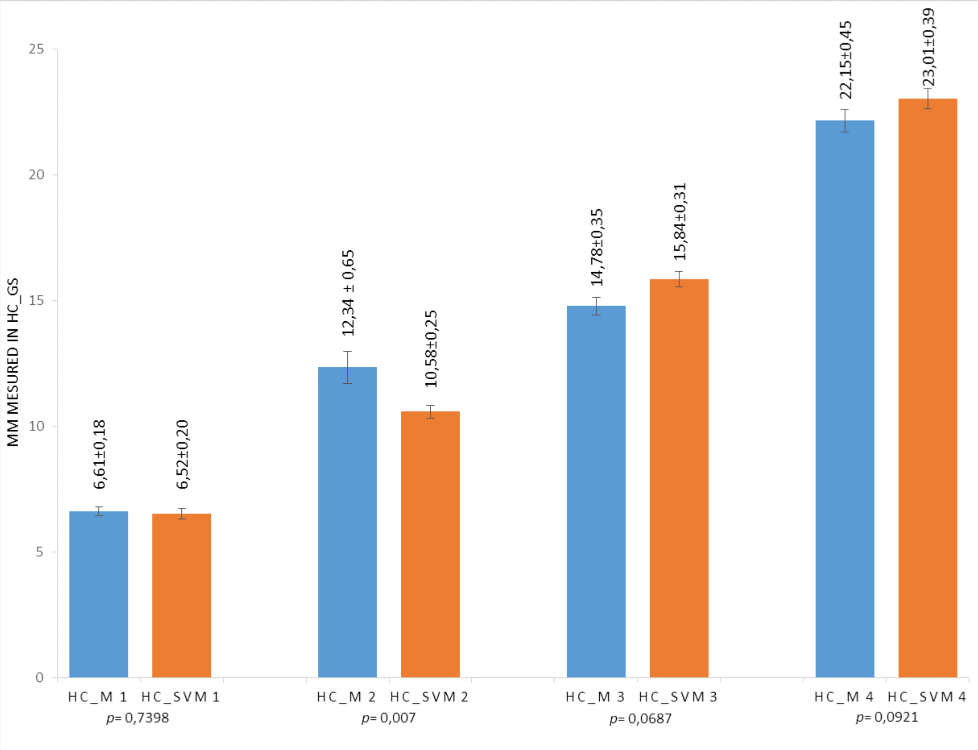

Fig. 2.5 Subcutaneous fat thickness of hams over muscle gluteus medius (mean ± SE mm) measured by HC_GS according to the category assigned by the classification system HC_M (left columns) and HC_SVM (right columns). At the bottom of columns the p-value of two-samples T-test between HC_M and H_SVM in each ham class.

Other factors may explain why HC_SVM obtained better results than HC_M. One of them is the probably operator fatigue, according to Font-i-Furnols et al. (2016) and Olsen et al. (2007), process repeatability or overexposure to a same class tends to lead to errors in classification (e.g. after classifying as HC4 a large number of hams, hams with slightly less fat tend to be classified into lower categories). Moreover, the cutting process of the carcasses usually obtain asymmetries in the two half-carcasses hindering the process of classification (Nissen et al., 2006). These factors could also affect the HC_GS measures. However, the HC_M measures were obtained on-line, with the speed of the production chain, while HC_GS was done out-line, without any speed chain. These factors suggest that using automatic algorithms with all information available and with a proper training phase it is possible to obtain better predictions of SFT of hams and therefore improve its classification.

Table 2-3 Percentage of correct, overestimate and underestimate classifications, between Manual prediction (HC_M) and automatic system (HC_SVM) from 400 hams according to breed and sex.

	HC_M			HC_SVM		
	Correct[a]	**Over[b]**	**Under[c]**	**Correct[a]**	**Over[b]**	**Under[c]**
SEX						
Female	63%	23%	14%	71%	15%	15%
Castrated males	73%	13%	15%	77%	12%	11%
Entire males	91%	2%	7%	89%	2%	9%
BREED						
(Large White x Landrace) x Duroc	69%	17%	14%	78%	14%	9%
(Large White x Landrace) x (Duroc x Landrace)	67%	19%	14%	79%	12%	9%
(Large White × Landrace) × Pietrain	71%	16%	13%	73%	11%	16%

[a] Correct, [b] Over Overestimate, [c] Under Underestimate

Table 2.3 shows the results of the classification of the two assessed methods. The results are shown in percentage and distributed according to successes, overestimated and underestimated categories. Those classifications are showed according to sex, at the top of the table, and according to breed, at the bottom of the table.

The results according to sex show that the HC_SVM obtained better results and the number of overestimate measures of castrated males and females was lower than in HC_M. Moreover, in both methods, the percentage of success in entire males was very high with a percentage of 91% in the case of HC_M and 89% in the case of HC_SVM. Those results suggest that better predictions are obtained due to entire males have leaner hams with low deviation of the SFT (see Table 2.2). Consequently, they were easy to predict as HC1.

Instead, regarding breed, HC_SVM obtained better results of overestimation and underestimation in all breeds except for crossbreed Pietrain underestimation, compared to HC_M. Therefore, correct predictions were similar between the two methods in the case of (Large White × Landrace) × Pietrain with

71% and 73% of correct predictions of HC_M and HC_SVM, respectively. Moreover, HC_SVM obtained better results when crossbreed included a Duroc line, compared to HC_M. Indeed correct predictions were a 9% and a 12% higher regarding the (Large White x Landrace) x Duroc and the (Large White x Landrace) x (Duroc x Landrace), respectively.

These results suggest that the HC_SVM better predicted ham classification than HC_M, especially in the fattest carcasses which mainly included carcasses from females, castrated males and Duroc crosses according to Gispert et al. (2010) and Wood, Enser, Whittington, Moncrieff, & Kempster (1989). This makes sense since HC_SVM improves mainly in the fatter groups of classification (HC2 to HC4). In contrast, the results obtained from the leaner carcasses are similar between HC_SVM and HC_M, which mainly included entire males and Pietrain crosses. Those results are especially relevant from an economic point of view, as the fattest hams are the ones that can be more valued for drying purposes.

2.3.2. Re-training of the algorithm with HC_GS

To assess whether it was possible to improve the HC_SVM algorithm, the SVM algorithm was re-trained using HC_GS as a response variable. As a result of this test, a model (HC_SVM2) was obtained with an accuracy of 75% and a coefficient k, of 0.67 (substantial agreement) was obtained. The results obtained in the HC_SVM2 model do not improve the percentage of success obtained in the classification of HC_SVM obtained in this study. This result suggests that the original SVM algorithm obtained by Masferrer et al (2018) was as good as the re-trained model HC_SVM2 probable because a large amount of data was used in HC_SVM. Although HC_SVM2 did not improve on the previous ones, it might be interesting for future work to explore other methods using HC_GS as a continuous variable.

2.3.3. Slaughterhouse profit

The impact of the correct hams classification is shown in Table 2.4. This table compares the potential benefits of a commercial slaughterhouse in 2017 classifying hams with a manual classification (HC_M) and with an automatic classification using an SVM model (HC_SVM). In order to quantify this potential benefit, an increase in price according to ham category was assigned as described in section 2.3.

Table 2-4 Comparison of the potential benefits of a year between classifying hams with the manual classification (HC_M) and with the automatic classification using an SVM model (HC_SVM). Data showed in the table is obtained from 2017 and belongs to a commercial slaughterhouse.

HC	Categ. (%)	Pigs	HC_M Correct	HC_M Profit increase	HC_SVM Correct	HC_SVM Profit increase
HC1	54	269400	237072	0 €	220908	0 €
HC2	28	138850	55540	42.038-140.127 €	83310	63.057-210.191€
HC3	13	66750	49395	149.548-249.247 €	52065	157.632-262.720 €
HC4	5	25000	19250	72.852-145.703 €	20250	76.636-153.272 €
Total	100	500000	361257	**264.438-535.078€**	376533	**297.325-626.183 €**

As expected, as showed in Table 2.4 better economic results are obtained using HC_SVM than using HC_M. Indeed, there is a difference between 30.000-90.000€, which represents an increase between 12-17% of the benefits.

Analysing these results according to HC, the HC_SVM obtained a potential benefit of 5% higher than using the HC_M in categories HC3 and HC4. Moreover, regarding the hams classified in HC2 category this difference was more than 50%. As expected, these results are due to the difference in accuracy of prediction in HC2 category between the two classification methods, being the HC_SVM better than the HC_M, but also because represents the 28% of all hams.

Furthermore, although it has not been taken into account in the previous economic analysis, the use of the SVM model allows to classify without an operator, saving the costs of production line personnel.

2.4. Conclusions

The results of the present study suggest that the use of automatic pattern recognition algorithms, and in this particular case, the SVM algorithm improves the prediction of the SFT measure and, therefore, the classification of the ham compared to HC_M. In addition, this method could allow the replacement of an operator in the production line, saving personnel costs, allowing faster chain speeds and reducing errors due to the fatigue of the operator. Moreover, it could improve subsequent processes in the cutting line, reducing the number of reprocessed hams and homogenizing batches for dry-curing processes. Consequently, this automatic HC_SVM method is more accurate and economically more beneficial for the meat industry than the manual HC_M method.

Acknowledgements

"Financial support from the Agency for Management of University and Research Grants (AGAUR) of the Catalan Government to Gerard Masferrer (2015 DI 00011) is grateful". CERCA Programme (Generalitat de Catalunya) is also acknowledged. The present study was supported by Grup Operatiu (Millora integral de la carn de porc a través d'estratègies productives a nivell de granja i eines innovadores, NIRS on-line per la seva classificació, a nivell d'escorxador. 562104820163A.

References

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning. Pattern Recognition* (Vol. 4). http://doi.org/10.1117/1.2819119

Bosi, P., Russo, V., & Paolo, B. (2004). The production of the heavy pig for high quality processed products. *J.ANIM.SCI*, *3*(21), 309–32145. http://doi.org/10.4081/ijas.2004.309

Brøndum, J., Egebo, M., Agerskov, C., & Busk, H. (1998). On-Line Pork Carcass Grading with the Autofom Ultrasound System. *Journal of Animal Science*, *76*(7), 1859–1868.

Buscailhon, S., Gandemer, G., & Monin, G. (1994). Time-related changes in intramuscular lipids of French dry-cured ham. *Meat Science*, *37*(2), 245–255. http://doi.org/10.1016/0309-1740(94)90084-1

Candek-Potokar, M., & Skrlep, M. (2012). Factors in pig production that impact the quality of dry-cured ham: a review. *Animal : An International Journal of Animal Bioscience*, *6*(2), 327–38. http://doi.org/10.1017/S1751731111001625

Cisneros, F., Ellis, M., McKeith, F. K., McCaw, J., & Fernando, R. L. (1996). Influence of Slaughter Weight on Growth and Carcass Characteristics, Commercial Cutting and Curing Yields, and Meat Quality of Barrows and Gilts from Two Genotypes. *Journal of Animal Science*, *74*(5), 925–933. http://doi.org/1996.745925

Font-i-Furnols, M., Čandek-Potokar, M., Daumas, G., Gispert, M., Judas, M., & Seynaeve, M. (2016). Comparison of national ZP equations for lean meat percentage assessment in SEUROP pig classification. *Meat Science*, *113*, 1–8. http://doi.org/10.1016/j.meatsci.2015.11.004

Font i Furnols, M., & Gispert, M. (2009). Comparison of different devices for predicting the lean meat percentage of pig carcasses. *Meat Science*, *83*(3), 443–446. http://doi.org/10.1016/j.meatsci.2009.06.018

Gispert, M., Àngels Oliver, M., Velarde, A., Suarez, P., Pérez, J., & Font i Furnols, M. (2010). Carcass and meat quality characteristics of immunocastrated male, surgically castrated male, entire male and female pigs. *Meat Science*, *85*(4), 664–670. http://doi.org/10.1016/j.meatsci.2010.03.021

Gispert, M., Font i Furnols, M., Gil, M., Velarde, A., Diestre, A., Carrión, D., … Plastow, G. S. (2007). Relationships between carcass quality parameters and genetic types. *Meat Science*, *77*(3), 397–404. http://doi.org/10.1016/j.meatsci.2007.04.006

Jain, A. K., Duin, P. W., & Jianchang Mao. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 4–37. http://doi.org/10.1109/34.824819

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. http://doi.org/10.2307/2529310

Marriott, N. G., Graham, P. P., & Claus, J. R. (1992). ACCELERATED DRY CURING OF PORK LEGS (HAMS): A REVIEW. *Journal of Muscle Foods*, *3*(2), 159–168. http://doi.org/10.1111/j.1745-4573.1992.tb00472.x

Masferrer, G., Carreras, R., Font-i-Furnols, M., Gispert, M., Marti-Puig, P., & Serra, M. (2018). On-line Ham Grading using pattern recognition models based on available data in commercial pig slaughterhouses. *Meat Science*, *143*, 39–45. http://doi.org/10.1016/j.meatsci.2018.04.011

Nissen, P. M., Busk, H., Oksama, M., Seynaeve, M., Gispert, M., Walstra, P., … Olsen, E. (2006). The estimated accuracy of the EU reference dissection method for pig carcass classification. *Meat Science*, *73*(1), 22–28. http://doi.org/10.1016/J.MEATSCI.2005.10.009

Olsen, E. V., Candek-Potokar, M., Oksama, M., Kien, S., Lisiak, D., & Busk, H. (2007). On-line measurements in pig carcass classification: Repeatability and variation caused by the operator and the copy of instrument. *Meat Science*, *75*(1), 29–38. http://doi.org/10.1016/j.meatsci.2006.06.011

Pulkrábek, J., Pavlík, J., & Valis, L. (2006). Pig carcass quality in relation to carcass lean meat proportion. *Czech Journal of Animal*. Retrieved from http://agriculturejournals.cz/publicFiles/52250.pdf

Ruiz-Carrascal, J., Ventanas, J., Cava, R., Andrés, A. I., & García, C. (2000). Texture and appearance of dry cured ham as affected by fat content and fatty acid composition. *Food Research International*, *33*(2), 91–95. http://doi.org/10.1016/S0963-9969(99)00153-2

Škrlep, M., Čandek-Potokar, M., Lukač, N. B., Povše, M. P., Pugliese, C., Labussière, E., & Flores, M. (2016). Comparison of entire male and immunocastrated pigs for dry-cured ham production under two salting regimes. *Meat Science*, *111*, 27–37. http://doi.org/10.1016/j.meatsci.2015.08.010

St. John, L. C., Young, C. R., Knabe, D. A., Thompson, L. D., Schelling, G. T., Grundy, S. M., & Smith, S. B. (1987). Fatty Acid Profiles and Sensory and Carcass Traits of Tissues from Steers and Swine Fed an Elevated Monounsaturated Fat Diet. *Journal of Animal Science*, *64*(5), 1441–1447. http://doi.org/10.2527/jas1987.6451441x

Toldrá, F., & Flores, M. (1998). The Role of Muscle Proteases and Lipases in Flavor Development During the Processing of Dry-Cured Ham. *Critical Reviews in Food Science and Nutrition*, *38*(4), 331–352. http://doi.org/10.1080/10408699891274237

Toldrá, F., Flores, M., & Sanz, Y. (1997). Dry-cured ham flavour: Enzymatic generation and process influence. *Food Chemistry*, *59*(4), 523–530. http://doi.org/10.1016/S0308-8146(97)00013-7

Warnants, N., Van Oeckel, M. J., & Boucqué, C. V. (1996). Incorporation of dietary polyunsaturated fatty acids in pork tissues and its implications for the quality of the end products. *Meat Science*, *44*(1–2), 125–144. http://doi.org/10.1016/S0309-1740(96)00029-0

Wood, J. D., Enser, M., Whittington, F. M., Moncrieff, C. B., & Kempster, A. J. (1989). Wood et al., 1989 Backfat composition in pigs Differences between fat thickness groups and sexes..pdf. *Livestock Production Science Elsevier Science Publishers B.V.*

3. SORTING HAMS USING BAGGED DECISION TREES IN A COMMERCIAL PIG SLAUGHTERHOUSE

Sorting Hams using Bagged Decision Trees in a commercial pig slaughterhouse

Gerard MASFERRER[a,b] Ricard CARRERAS[b], Pere MARTI-PUIG [a] and Moises SERRA [a]

[a] *Information and Digital Technology Department, UVic-UCC, Vic (Barcelona)*
[b] *Mafrica.SA, Paratge Can Canals Nou, S/N 08250 Sant Joan de Vilatorrada (Barcelona)*

Abstract

Classification of pig carcasses according to its characteristics is a determining factor in slaughterhouses, since allows the optimization of production and improves the performance in cutting plants and in other subsequent processes. Usual criteria in carcasses classification are the weight and the fat content of the ham, especially in regions such as Spain where a significant proportion of hams are used for curing to produce Jamón Serrano. The objective of this study is to compare different models based on Decision Trees and using intrinsic data of pigs to classify hams in four groups according to their predicted weight. The model presented is based on Bagged Decision Trees and use as input: the weight and the lean meat percentage (LMP) of the carcass, the sex and the breed of the pig and the manual classification of the ham according to the thickness of the subcutaneous fat. The results show a success rate of 81.7%, improving by 4.4% the results obtained with a more straightforward decision tree based only on the weight and the LMP.

Keywords. dry-cured hams; ham classification; pattern recognition, decision trees

3.1. Introduction

The ham is one of the most valuable pieces of pork, furthermore represents a significant percentage of the carcass, between 25-30% [1,2], and is especially appreciated in countries such as Spain or Italy where products such as Serrano Ham are made by subjecting the ham to curing processes.

The curing process involves many factors that define the final quality of the product [3]. The weight of the ham is one of them as it determines different aspects such as the amount of salt in the product or the curing time [2,4]. Therefore, hams and carcasses are usually grouped in slaughterhouses according to customer specifications or their subsequent production process.

Different factors affect the weight of the ham. As expected, it is mainly determined by the weight of the animal but with significant variations according to the conformation of the ham, the subcutaneous fat, and consequently the percentage of lean meat. On the other hand, all these factors are closely related to breed and sex, for example, Pietran crossbreds are leaner than Duroc crossbreds [5], and females tend to deposit more subcutaneous fat than entire males [6].

Although the early determination of the weight of the ham in the line is of great importance, in some slaughterhouses there is no pre-classification, and the classification is carried out in the cutting room at the end of the process with the finished product. In these cases, all hams are processed equally as differentiated processes according to ham weight are not allowed.

There are different methods to determine the weight of the ham. In the most straightforward cases, the weight of the carcass and its lean meat percentage (LMP), a parameter that must be determined compulsorily according to the (EU) 2017/1182, are used to classify hams according to these characteristics. In other slaughterhouses, automatic or semi-automatic systems are used to determine LMP [7], which in its most advanced and expensive versions predict the weight of the primary cuts, among them, the ham. However, it is interesting that all slaughterhouses, having or not advanced versions of automatic and semi-automatic systems, can pre-classify hams according to its weight in a more precise way. Therefore, it would be interesting to find a method using only compulsory data for slaughterhouses (LMP) and intrinsic data of pigs (weight, sex, breed and manual measure of the thickness of subcutaneous fat).

Decision Trees are used successfully in different areas [8,9] and are easy to implement. There are different configurations when developing these algorithms, one of them is the Bagged Decision Trees, where *n* Decision Trees are combined. In these algorithms, the effect of overfitting is reduced, and generalization is improved. Moreover, those algorithms are more accurate in complex systems, that is, with more inputs [10].

This study aims to present and compare three models based on Decision Trees and using intrinsic pig data for the classification of carcasses according to the weight of the ham.

3.2. Material and Methods

3.2.1. Animals and facilities

This study was carried out with data obtained during April 2018 from pigs fattened in Spanish commercial slaughterhouse (MAFRICA S.A.) located in Sant Joan de Vilatorrada, Catalonia. Pigs were slaughtered. After scalding, they were tagged with two radio-frequency identifiers (RFID), one for each half-carcass, and were scanned using the ultrasound AutoFom-III system [11]. Then pigs were eviscerated and split according to standard commercial procedures using an automatic robotic system. After that, the two half-carcasses were weighted and an experimented operator visually determined the sex of the pig (female, entire male or castrated male). Then the left half carcass was manually classified according to the subcutaneous fat thickness (SFT) in four categories: (thin (0-10mm), standard (11-15 mm), semi-fat (16-20 mm) and fat (>20 mm)), as described in [12]. Approximately, after 30 min carcasses were processed in the cutting room where the head, the loin, the ribs and the shoulders were removed. The rest of the carcass, including the ham, the belly, and the neck, was stored and ordered in cooling during approximately 24h.

3.2.2. Hams processing

After 24h in the cooling chamber, these carcasses were at approximately 4°C. The ham was removed from the rest of the carcass, was classified according to Table 3.1 and was processed to give the final shape. At this point, hams were weighed individually with an in-line scale (see Figure 3.1). By using an RFID identifier, weight data was stored in the database of the corresponding carcass.

a) b)

Fig. 3.1 a) Processed ham. b) In-line scale to weigh hams and carcasses.

For this study the weight of 13,813 hams, corresponding to a week's production were recorded. Hams were classified according to the weight categories showed in Table 3.1.

Table 3-1 Ham classification according to its weight.

Class	Weight Range (Kg)
S	<12
M	Between 12-14
L	Between 14-16
XL	>16

3.2.3. Statistical analysis

All data collected in the production line was stored in a database including the following parameters per sample: LMP (lean meat percentage); SEX ((1) females, (2) entire males and (3) castrated males); WGT (warm carcass weight); BREED (there are three categories expressed from the breed lines of the mother and father: (1) (Large White x Landrace) x Pietrain, (2) (Large White x Landrace) x Duroc and (3) (Large White x Landrace) x (Duroc x Landrace)); HC (manually assigned class according to subcutaneous fat thickness of the ham, with a pattern according to the procedure described in [2]). The elements of this database were used as inputs to predict the weight category to which the ham belongs according to the categories established in Table 3.1, where the category (S, M, L, XL) is the response to predict.

Three models were tested, (M1) a Decision Tree using WGT and LMP features as input, (M2) Decision Tree using as input, WGT, LMP, SEX, BREED and HC, (M3) a model using Bagged Trees with WGT, LMP, SEX, BREED and HC.

The use of Bagging predictors is a common method to aggregate multiple versions of more simple predictors. The aggregation averages the outcomes of single predictors by voting their prediction results. That strategy usually improves final accuracy and especially provides overfitting protection [10]. The models were trained and validated using 5-Fold Cross-validation [13] and, after that, were evaluated according to their accuracy. Also, with the aim of evaluating the categories of hams individually, three confusion matrices were represented using a cross-table with the real values, obtained when the hams were weighed, and the values predicted with the different models. MATLAB and Signal Processing Toolbox™ (Matlab R2018a; The MathWorks, Inc, 1988–2018) have been used to develop and test all the models and algorithms.

3.3. Results and discussion

Figure 3.2 shows the confusion matrix presented as a cross-table where the percentage of success according to the predicted class and the real class, obtained when the ham was weighted, are presented. Model 1 (M1) achieves a total accuracy of 77.3% (in the prediction of the four classes). This is a good result since only the carcass weight and LMP parameter were used, but foreseeable considering that the ham represents between 25-30% of the carcass according to [1,2] and the high correlation between LMP and SFT [14] and the difference between the weight of the lean meat and the weight of the fat.

When comparing the different ham categories, it could be observed that the central ones, i.e. M and L, obtain a higher correct classification ratio (83% and 80%, respectively) than the extreme ones, i.e. S and XL (59% and 66%, respectively). These results suggest that the fewer number of hams (S=873, M=5147, L=5568, XL=2225) in the extreme categories compared with the central ones could influence on the correct classification ratio since the Decision Tree tends to underestimate extreme classes.

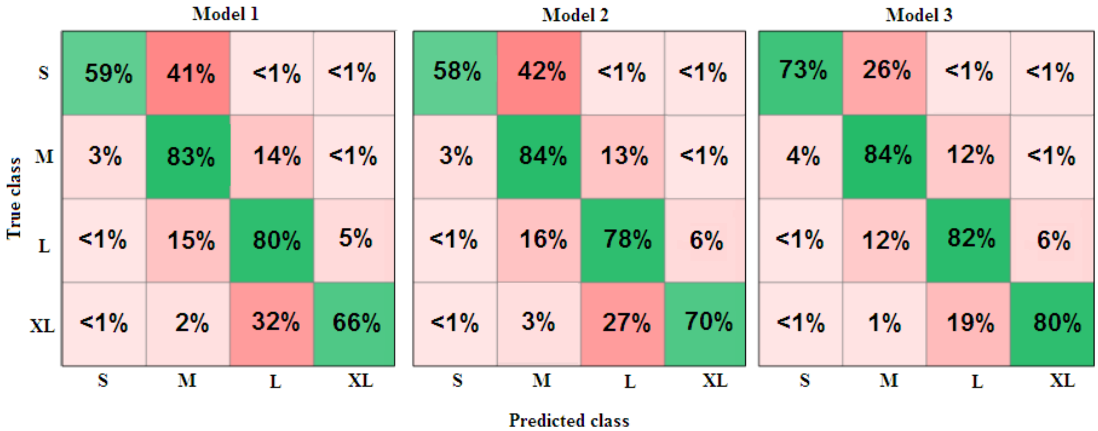

Fig. 1.2 Confusion matrix of the three models used in this study. On the top left Model 1(Decision Tree with WGT and Lean meat percentage (LMP)) is shown; on the center Model 2 (Decision Tree with WGT, LMP, SEX, BREED and HC) is shown and, at the right, Model 3 (Bagged Tree using same Model 2 inputs) is shown.

In model 2 (M2), where SEX, BREED and HC are incorporated, achieves the success of 77.7%, improving 0.4% compared to M1. This improvement can be explained by the incorporation of SEX and BREED, which probably improves the estimation of fat carcass [1,6,15] and, especially, by the incorporation of HC which provides direct information of SFT measure.

Finally, model 3 (M3), use the same inputs as M2 and Bagged Trees, with our particular configuration, that is, with a bag of 30 Decision Trees. In this case M3 achieves the success of 81.7%, improving by 4% compared to M2. Comparing the different ham categories, results show an improvement by 4% in the classification of the central category L. Furthermore, the accuracy of extreme groups, S and XL, improves a 15% and 10%, respectively compared to M2. Bagged Trees, slightly increases the complexity of the model, but in contrast, improves the classification, particularly in the extreme classes that are less represented in our dataset [10].

3.4. Conclusions

The use of Bagged Decision Trees and intrinsic data of pigs available in all slaughterhouses may lead to an improvement in ham weight classification systems. Only using this improvement allows better carcasses and sub-products classification according to customer specifications. Moreover, the action of grouping carcasses in a more optimal way improves the production line and linearized processes. This model is particularly interesting in slaughterhouses that only use measures that can be obtained without complex systems to scan carcasses such as Autofom (AutoFom III, VCS 2000) [9].

Acknowledgements

Aquest treball ha estat realitzat en el marc del programa de doctorat en Ciències Experimentals i Tecnologies de la UVic-UCC, amb l'ajut de Doctorat Industrial concedit per l'Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) a Gerard Masferrer.

References

[1] M. Gispert, M. Font i Furnols, M. Gil, A. Velarde, A. Diestre, D. Carrión, A.A. Sosnicki, and G.S. Plastow, Relationships between carcass quality parameters and genetic types, *Meat Sci.* **77** (2007) 397–404.

[2] F. Cisneros, M. Ellis, F.K. McKeith, J. McCaw, and R.L. Fernando, Influence of Slaughter Weight on Growth and Carcass Characteristics, Commercial Cutting and Curing Yields, and Meat Quality of Barrows and Gilts from Two Genotypes, *J. Anim. Sci.* **74** (1996) 925–933. doi:1996.745925.

[3] M. Candek-Potokar, and M. Skrlep, Factors in pig production that impact the quality of dry-cured ham: a review., *Animal.* **6** (2012) 327–38. doi:10.1017/S1751731111001625.

[4] M.A. Latorre, G. Ripoll, E. García-Belenguer, and L. Ariño, The increase of slaughter weight in gilts as a strategy to optimize the production of Spanish high quality cured ham, *J.Anim. Sci.* **87** (2009) 1464–1471.

[5] D.B. Edwards, R.O. Bates, and W.N. Osburn, Evaluation of Duroc- vs. Pietrain-sired pigs for carcass and meat quality measures, *J. Anim. Sci.* **81** (2003) 1895. doi:10.2527/2003.8181895x.

[6] M. Gispert, M. Àngels Oliver, A. Velarde, P. Suarez, J. Pérez, and M. Font i Furnols, Carcass and meat quality characteristics of immunocastrated male, surgically castrated male, entire male and female pigs, *Meat Sci.* **85** (2010) 664–670. doi:10.1016/j.meatsci.2010.03.021.

[7] M. Font i Furnols, and M. Gispert, Comparison of different devices for predicting the lean meat percentage of pig carcasses, *Meat Sci.* **83** (2009) 443–446. doi:10.1016/j.meatsci.2009.06.018.

[8] M.A. Friedl, and C.E. Brodley, Decision tree classification of land cover from remotely sensed data, *Remote Sens. Environ.* **61** (1997) 399–409. doi:10.1016/S0034-4257(97)00049-7.

[9] S.R. Safavian, and D. Landgrebe, A Survey of Decsion Tree Clasifier Methodology, *IEEE Trans. Syst. Man. Cybern.* **21** (1990) 660–674.

[10] L. Breiman, Bagging predictors, *Mach. Learn.* **24** (1996) 123–140. doi:10.1007/BF00058655.

[11] J. Brøndum, M. Egebo, C. Agerskov, and H. Busk, On-Line Pork Carcass Grading with the Autofom Ultrasound System, *J. Anim. Sci.* **76** (1998) 1859–1868.

[12] G. Masferrer, R. Carreras, M. Font-i-Furnols, M. Gispert, P. Marti-Puig, and M. Serra, On-line Ham Grading using pattern recognition models based on available data in commercial pig slaughterhouses, *Meat Sci.* **143** (2018) 39–45. doi:10.1016/j.meatsci.2018.04.011.

[13] C.M. Bishop, Pattern Recognition and Machine Learning, 2006. doi:10.1117/1.2819119.

[14] J.D. Wood, M. Enser, F.M. Whittington, C.B. Moncrieff, and A.J. Kempster, Wood et al., 1989 Backfat composition in pigs Differences between fat thickness groups and sexes..pdf, *Livest. Prod. Sci. Elsevier Sci. Publ. B.V.* **22** (1989) 351–362.

[15] M. Font-i-Furnols, M. Gispert, J. Soler, M. Diaz, J.A. Garcia-Regueiro, I. Diaz, and M.C. Pearce, Effect of vaccination against gonadotrophin-releasing factor on growth performance, carcass, meat and fat quality of male Duroc pigs for dry-cured ham production, *Meat Sci.* **91** (2012) 148–154. doi:10.1016/j.meatsci.2012.01.00.00000aacb35d&acdnat=1493822848_b5a764336449f25b0d2b04164c239dad

III. DISCUSSIÓ GENERAL

III.I L'ús de tècniques de reconeixement de patrons i classificació automàtica en escorxadors

El sector porcí disposa d’un volum molt important de dades degut, principalment, a la producció mundial de carn de porc i conseqüentment la quantitat de caps sacrificats anualment. A més a més, la inclusió de sistemes de classificació de canals en les cadenes de producció fa que, per una simple qüestió aritmètica, es disposi de moltes dades estadísticament significatives. Els estudis realitzats en aquesta tesi doctoral posen de manifest que l'ús d'aquestes dades en tècniques de processat, àmpliament utilitzades en altres sectors (Bishop, 2006), tenen un gran potencial a l'hora d'optimitzar i millorar procesos industrials de classificació en línia.

En aquest sentit, en els diferents estudis s’observa una millora en les classificacions a mesura que s’incorporen paràmetres procedents d’equips automàtics i de mesures preses manualment per opearis a la línia. Aquesta millora és molt variable en funció del paràmetre incorporat, tal i com s’observa en el capítol 1. Així doncs, en alguns casos la millora en la classificació no és molt significativa al tractar-se de paràmetres molt correlacionats entre ells, com per exemple l'F34, l'F_GM1 i l'F_GM2. Relacionat amb la incorporació d'aquestes dades, cal puntualitzar que no s’ha tingut en compte l'error potencial que s'acumula pel fet d’utilitzar més paràmetres, ja que, tal i com es va comprovar en els treballs previs a l’escorxador, com més equips i més punts d'obtenció de dades siguin necessaris per classificar, més augmenta la probabilitat d’error de la predicció (del propi equip o del sistema RFID (Radio Frequency Identifier) per emmagatzemar la informació). En treballs futurs seria interessant estudiar l’error que s'atribueix a la incorporació de nous equips que generen nous paràmetres utilitzats en els algoritmes per a la classificació en línia.

En els tres capítols de la tesi observem una millora en els resultats obtinguts quan s’incorpora informació intrínseca dels animals (sexe, raça i pes) en els models de classificació. Per exemple, en el capítol 1 observem que la incorporació del pes de la canal, del sexe i de la raça millora el percentatge d'encert en molts dels models de classificació. Aquesta millora s'observa també en el capítol 3 on el segon model (M2) millora lleugerament quan incorporem els parametres sexe i raça. De manera similar en el capítol 3 s'observa que l'HC_SVM prediu millor la classificació del pernil que l'HC_M, especialment en les canals més grasses segons sexe (femelles, mascles castrats) i segons raça en creuaments de Duroc ((Large White x Landrace) x Duroc i (Large White x Landrace) x (Duroc x Landrace)). Aquests resultats coincideixen amb Gispert et al., 2007 i Wood et al., 1989, on s'observa que existeix una relació clara entre raça i greix subcutani en el pernil. També observem una relació directa entre el sexe dels animals i el greix subcutani del pernil (Font-i-Furnols et al., 2012; Gispert et al., 2010). Aquesta informació és fàcilment representable en arbres de decisió i incorporable en molts models.

En determinats països europeus, com per exemple Dinamarca (Hinrichsen, 2010), el nivell d'automatizcació als escorxadors és molt superior a la majoria dels escorxadors d'Espanya, on encara segueixen un procés de producció molt manual. És per aquest motiu que en aquests escorxadors els operaris encara duen a terme un conjunt de tasques laborioses i repetitives, tot i que lentament s'hi van introduint noves tecnologies. En aquest sentit cal destacar que tant en el capítol 1 com en el capítol 2 s'observa un possible efecte de la fatiga en la classificació realitzada per un operari (Font-i-Furnols et al., 2016; Olsen et al., 2007). En el capítol 1 s'utilitza la classificació manual com a mètode de referència, per tant, tot i que s'obté un model basat en SVM que imita a l'operari en un 73%, part de classificació errònia podria ser explicada per errors humans. En efecte, en el capítol 2 s'observa que la classificació automàtica HC_SVM té un percentatge d'encert del 75.3%, comparat amb el 69.8% de la classificació manual (HC_M). Això pot explicar, en part, la millora obtinguda utilitzant algoritmes automàtics per classificar, en els què s'elimina l'efecte de la fatiga i de la subjectivitat en la classificació.

Per altra banda, en els diferents estudis realitzats s'observa una reducció de pernils classificats en categories no adjacents quan s'utilitzen dades proporcionades per aparells en comptes de dades proporcionades per operaris. En el capítol 1 el millor dels models obtinguts (SVM Medium Gaussian) imita la classificació manual amb un 73% d'encert, amb un 3,6% dels pernils classificats erròniament a categories no adjacents. Aquest valor és molt similar als obtinguts al capítol 2 en la classificació manual (HC_M) amb un 3,8% respecte la classificació "Golden Standard" (HC_GS) de referència, mentre que en el cas de la classificació realitzada amb el model de SVM (HC_SVM) l'error a categories no adjacents es veu reduïda a un 0,6%. Aquests resultats suggereixen, tal com s'apunta en el capítol 2, que la classificació automàtica presenta menys desviació en les categories i conseqüentment millora la classificació i els errors a categories no adjacents respecte a la classificació manual.

Cal destacar que a diferència de sectors on l'ús de tècniques de reconeixement de patrons i processat de dades és àmpliament utilitzat des de fa molts anys, en el sector carni hi ha un gran potencial de millora. En el conjunt de treballs publicats en aquesta tesi s'han aplicat tècniques àmpliament conegudes i aplicades en altres sectors amb la finalitat específica de classificar el pernil i s'han obtingut resultats satisfactoris. Els casos estudiats en les diferents publicacions estan basats en el pernil degut a la importància que té aquesta peça en el sector porcí i, específicament, en el procés productiu de Mafrica, l'escorxador on s'ha dut a terme el doctorat industrial.

En efecte, en els tres capítols on s'utilitzen mètodes de reconeixement de patrons per a classificar pernil suggereixen una millora econòmica tant per l'escorxador com pel client final. Així doncs, tot i ser un valor molt variable segons l'escorxador, el producte, etc., en el capítol 2 s’estima que l’ús de l’algoritme automàtic podria suposar un increment d’entre un 12% i un 17% del marge de benefici, tot gràcies a la classificació. A més a més, en els diferents capítols es conclou en apuntar una millora gràcies a la disminució de pernils que s’han de reprocessar, així com l'optimització dels processos en cadena i l’obtenció de lots de producte més homogenis.

Aquest conjunt de tècniques hauria de ser fàcilment extrapolable a altres parts del porc. Per tant l'aplicació de les tècniques presentades en aquesta tesi només és una primera pinzellada del potencial real que poden arribar a tenir en el sector. Cal destacar també la necessitat de transferència de coneixement del sector TIC (tecnologies de la informació i la comunicació) cap a aquesta indústria.

I.V. CONCLUSIONS

Dels diferents treballs publicats se n'extreuen les conclusions següents:

- L'ús de tècniques de reconeixement de patrons utilitzades amb dades disponibles en els escorxadors ens permet classificar pernils segons l'espessor de greix subcutani.
- L'ús d'algoritmes de reconeixement automàtic de patrons, concretament l'algoritme SVM amb nucli Gaussià, classifica millor els pernils segons el seu espessor de greix subcutani comparat amb una classificació manual.
- Els models de reconeixement de patrons poden contribuir a millorar el rendiment dels escorxadors, ja que permeten optimitzar processos en les sales d'especejament, augmentar la velocitat de classificació i probablement reduir errors humans produïts per fatiga.
- L'aplicació de models de reconeixement de patrons per a la classificació automàtica de pernils pot incrementar el benefici econòmic degut a la millora de la classificació.
- La incorporació de dades intrínseques dels animals (raça i sexe) en arbres de decisió permet millorar l'agrupació de pernils per pes en comparació a sistemes on només s'utilitza el pes de la canal i el seu percentatge de magre.
- Els algoritmes de reconeixement de patrons contribueixen en millorar la creació de lots tant per pes com per greix subcutani. Aquesta millora no només és beneficiosa pels processos productius de l'escorxador i de la sala d'especejament, si no que també pot tenir un impacte econòmic sobre el procés que durà a terme el client (com per exemple l'avantatge que suposa la homogeneïtat de lots en processos de curat i/o cuit).

V. BIBLIOGRAFIA GENERAL

V.I .Referències

Balakrishnama, S., & Ganapathiraju, A. (1998). Linear Discriminant Analysis a Brief Tutorial. Compute, 11, 1–9. http://www.isip.piconepress.com/publications/reports/1998/isip/lda/

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Pattern Recognition (Vol. 4). http://doi.org/10.1117/1.2819119

Bosi, P., Russo, V., & Paolo, B. (2004). The production of the heavy pig for high quality processed products. J.ANIM.SCI, 3(21), 309–32145. http://doi.org/10.4081/ijas.2004.309

Breiman, l. Bagging predictors, Mach. Learn. 24 (1996) 123–140. doi:10.1007/BF00058655.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2(2), 121–167. http://doi.org/10.1023/A:1009715923555

Candek-Potokar, M., & Skrlep, M. (2012). Factors in pig production that impact the quality of dry-cured ham: a review. Animal : An International Journal of Animal Bioscience, 6(2), 327–38. http://doi.org/10.1017/S1751731111001625

Carabús, A., Gispert, M., Brun, A., Rodríguez, P. & Font-i-Furnols, M. (2014). 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. http://doi.org/10.1016/j.livsci.2014.10.005

Cilla, I., Altarriba, J., Guerrero, L., Gispert, M., Martínez, L., Moreno, C., Roncalés, P. (2006). Effect of different Duroc line sires on carcass composition, meat quality and dry-cured ham acceptability. Meat Science, 72(2), 252–260. http://doi.org/10.1016/j.meatsci.2005.07.010

Cisneros, F., Ellis, M., McKeith, F. K., McCaw, J., & Fernando, R. L. (1996). Influence of Slaughter Weight on Growth and Carcass Characteristics, Commercial Cutting and Curing Yields, and Meat Quality of Barrows and Gilts from Two Genotypes. *Journal of Animal Science*, *74*(5), 925–933. http://doi.org/1996.745925

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 4.7.17, L171/74-L171-99.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. http://doi.org/10.1109/TIT.1967.1053964

Daumas, G. (2011). Non-electronic techniques to classify pig carcasses in small slaughterhouses. In Second International Virtual Conference on pork Quality. Novembere, 05to Decembere (p. p 06-2001).

Dirinck, P., Van Opstaele, F., & Vandendriessche, F. (1997). Flavour differences between northern and southern European cured hams. Food Chemistry, 59(4), 511–521. http://doi.org/10.1016/S0308-8146(97)00012-5

Edwards, D. B., Bates, R. O., & Osburn, W. N. (2003). Evaluation of Duroc- vs. Pietrain-sired pigs for carcass and meat quality measures. Journal of Animal Science, 81(8), 1895. http://doi.org/10.2527/2003.8181895x

Eurostat 2018.- Database (n.d.). Retrieved April 15, 2019, from https://ec.europa.eu/eurostat/data/database

Fàbrega, E., Gispert, M., Tibau, J., Hortós, M., Oliver, M. A., & Font i Furnols, M. (2011). Effect of housing system, slaughter weight and slaughter strategy on carcass and meat quality, sex organ development and androstenone and skatole levels in Duroc finished entire male pigs. Meat Science, 89(4), 434–439. http://doi.org/10.1016/j.meatsci.2011.05.009

FAOSTAT, 2017. Food and Agriculture Organization of the United Nations. (n.d.). Retrieved April 30, 2019, from http://www.fao.org/faostat/en/#data

Fisher, R. A. (1936). The use of multiple measurements in taxanomic problems. Annals of Eugenics, 7(2), 179–188. http://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Font-i-Furnols, M., Čandek-Potokar, M., Daumas, G., Gispert, M., Judas, M., & Seynaeve, M. (2016). Comparison of national ZP equations for lean meat percentage assessment in SEUROP pig classification. Meat Science, 113, 1–8. http://doi.org/10.1016/j.meatsci.2015.11.004

Font-i-Furnols, M., Gispert, M., Soler, J., Diaz, M., Garcia-Regueiro, J. A., Diaz, I., & Pearce, M. C. (2012). Effect of vaccination against gonadotrophin-releasing factor on growth performance, carcass, meat and fat quality of male Duroc pigs for dry-cured ham production. Meat Science, 91(2), 148–154. http://doi.org/10.1016/j.meatsci.2012.01.008

Font i Furnols, M., & Gispert, M. (2009). Comparison of different devices for predicting the lean meat percentage of pig carcasses. Meat Science, 83(3), 443–446. http://doi.org/10.1016/j.meatsci.2009.06.018

Friedl, M.A.and Brodley, C.E. Decision tree classification of land cover from remotely sensed data, Remote Sens. Environ. 61 (1997) 399–409. doi:10.1016/S0034-4257(97)00049-7.

Gispert, M., Àngels Oliver, M., Velarde, A., Suarez, P., Pérez, J., & Font i Furnols, M. (2010). Carcass and meat quality characteristics of immunocastrated male, surgically castrated male, entire male and female pigs. Meat Science, 85(4), 664–670. http://doi.org/10.1016/j.meatsci.2010.03.021

Gispert, M., Font i Furnols, M., Gil, M., Velarde, A., Diestre, A., Carrión, D., … Plastow, G. S. (2007). Relationships between carcass quality parameters and genetic types. Meat Science, 77(3), 397–404. http://doi.org/10.1016/j.meatsci.2007.04.006

Hinrichsen, L. (2010). Manufacturing technology in the Danish pig slaughter industry. *Meat Science*, *84*(2), 271–275. http://doi.org/10.1016/J.MEATSCI.2009.03.012

Kempster, A. J., Chadwick, J. P., & Jones, D. W. (1985). An evaluation of the Hennessy grading probe and the SFK Fat-O-Meater for use in pig carcass classification and grading. Animal Production, 40(2), 323–329. http://doi.org/10.1017/S0003356100025447

Latorre, M. A., García-Belenguer, E., & Ariño, L. (2008). The effects of sex and slaughter weight on growth performance and carcass traits of pigs intended for dry-cured ham from Teruel (Spain). Journal of Animal Science, 86(8), 1933–1942. http://doi.org/10.2527/jas.2007-0764

Latorre, M.A. G. Ripoll, E. García-Belenguer, and L. Ariño, The increase of slaughter weight in gilts as a strategy to optimize the production of Spanish high quality cured ham, *J.Anim. Sci.* **87** (2009) 1464–1471

Lucas, D., Brun, A., Gispert, M., Carabús, A., Soler, J., Tibau, J., & Font-i-Furnols, M. (2017). Relationship between pig carcass characteristics measured in live pigs or carcasses with Piglog, Fat-o-Meat'er and computed tomography. Livestock Science, 197, 88–95. http://doi.org/10.1016/j.livsci.2017.01.010

MAPAMA 2017 Estadísticas agrarias: Ganadería. (n.d.). Retrieved April 15, 2019, from https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/ganaderia/

Nanni Costa, L., Stelletta, C., Cannizzo, C., Gianesella, M., Lo Fiego, D. P., & Morgante, M. (2010). The use of thermography on the slaughter-line for the assessment of pork and raw ham quality. Italian Journal of Animal Science, 6(1s). http://doi.org/10.4081/ijas.2007.1s.704

Olsen, E. V., Candek-Potokar, M., Oksama, M., Kien, S., Lisiak, D., & Busk, H. (2007). On-line measurements in pig carcass classification: Repeatability and variation caused by the operator and the copy of instrument. Meat Science, 75(1), 29–38. http://doi.org/10.1016/j.meatsci.2006.06.011

Realini, C. E., Duran-Montgé, P., Lizardo, R., Gispert, M., Oliver, M. A., & Esteve-Garcia, E. (2010). Effect of source of dietary fat on pig performance, carcass characteristics and carcass fat content, distribution and fatty acid composition. Meat Science, 85(4), 606–612. http://doi.org/10.1016/j.meatsci.2010.03.011

Safavian, S.R. and Landgrebe, D.A Survey of Decsion Tree Clasifier Methodology, IEEE Trans. Syst. Man. Cybern. 21 (1990) 660–674.

St. John, L. C., Young, C. R., Knabe, D. A., Thompson, L. D., Schelling, G. T., Grundy, S. M., & Smith, S. B. (1987). Fatty Acid Profiles and Sensory and Carcass Traits of Tissues from Steers and Swine Fed an Elevated Monounsaturated Fat Diet. *Journal of Animal Science*, *64*(5), 1441–1447. http://doi.org/10.2527/jas1987.6451441x

Toldrá, F., & Flores, M. (1998). The Role of Muscle Proteases and Lipases in Flavor Development During the Processing of Dry-Cured Ham. *Critical Reviews in Food Science and Nutrition*, *38*(4), 331–352. http://doi.org/10.1080/10408699891274237

Toldrá, F., Flores, M., & Sanz, Y. (1997). Dry-cured ham flavour: Enzymatic generation and process influence. *Food Chemistry*, *59*(4), 523–530. http://doi.org/10.1016/S0308-8146(97)00013-7

Tous, N., Lizardo, R., Vilà, B., Gispert, M., Font-I-Furnols, M., & Esteve-Garcia, E. (2014). Effect of reducing dietary protein and lysine on growth performance, carcass characteristics, intramuscular fat, and fatty acid profile of finishing barrows 1. J. Anim. Sci, 92, 129–140. http://doi.org/10.2527/jas2012-6222

Vapnik, V. N., & Chervonenkis, A. (1964). A note on one class of perceptrons. Automation and Remote Control, 25(6), 937–945. Retrieved from https://scholar.google.es/scholar?q=Vladimir+Vapnik+Alexey+Ya.+Chervonenkis&hl=ca&as_sdt=0%2C5&as_ylo=1960&as_yhi=1965

Warnants, N., Van Oeckel, M. J., & Boucqué, C. V. (1996). Incorporation of dietary polyunsaturated fatty acids in pork tissues and its implications for the quality of the end products. *Meat Science*, *44*(1–2), 125–144. http://doi.org/10.1016/S0309-1740(96)00029-0

Wood, J. D., Enser, M., Whittington, F. M., Moncrieff, C. B., & Kempster, A. J. (1989). Wood et al., 1989 Backfat composition in pigs Differences between fat thickness groups and sexes..pdf. *Livestock Production Science Elsevier Science Publishers B.V.* Retrieved from http://ac.els-cdn.com/0301622689900663/1-s2.0-0301622689900663-main.pdf?_tid=02bf6b6e-300f-11e7-a6c9-

Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Sheard, P. R., Richardson, R. I., … Whittington, F. M. (2008). Fat deposition, fatty acid composition and meat quality: A review. Meat Science. http://doi.org/10.1016/j.meatsci.2007.07.019

Wood, J. D., Nute, G. R., Richardson, R. I., Whittington, F. M., Southwood, O., Plastow, G., … Chang, K. C. (2004). Effects of breed, diet and muscle on fat deposition and eating quality in pigs. Meat Science, 67(4), 651–667. http://doi.org/10.1016/j.meatsci.2004.01.007