UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Department of animal medicine, production and healh (MAPS)

PhP course in: Scienze Veterinarie
XXXIV cycle

APPLICATION OF DEEP LEARNING NEURAL NETWORKS ON COMPUTED TOMOGRAPHY IMAGES OF CANINE HEPATIC AND SPLENIC MASSES

Coordinator: Ch.mo Prof. Mattia Cecchinato
Supervisor: Ch.mo Dr. Tommaso Banzato

PhD student: Silvia Burti

SUMMARY

CHAPTER ONE

Introdution

FOCAL LIVER LESIONS IN DOGS

Primary hepatic tumours are uncommon in dogs, counting for less than 1.5% of all the tumours in this species. Metastases are 2.5 times more frequent, especially from splenic, pancreatic, and gastroenteric neoplasia, that reach the liver through portal system [1,2].

Malignant primary hepatic tumours can be divided into 4 categories, based on their cytological origin: hepatocellular, biliary, neuroendocrine, and mesenchymal tumours. Furthermore, based on their macroscopic growth, we can distinguish: massive (when a single and big mass is present usually in a single hepatic lobe), nodular (when multifocal nodules in more than one hepatic lobe are present), or diffused (when the entire hepatic parenchyma is involved) tumours [1].

The prognosis is strictly related to the tumour histotype and morphology. Indeed, when a single lobe is interested by a massive tumour, even if malignant and with big dimensions, the prognosis is more favourable. Indeed, a complete surgical resection of the entire tumour is more probable and the risk of metastasis is lower [1]. On the contrary, when the tumour is characterised by having multifocal nodules or diffusive pattern, the life expectancy is lower [1].

Hepatocellular tumours include hepatocellular adenoma (or hepatoma) (HCA) and hepatocellular carcinoma (HCC).

HCA is the most common benign neoplasia in elderly dogs, and usually is an incidental finding during ultrasonographic, tomographic, or necroscopic examinations. HCA is not characterized by specific symptoms, even when the surrounding parenchyma is compression on. Furthermore, HCA does not represent a pre-neoplastic form of the hepatocellular carcinoma [3].

HCC represents the malignant primary hepatic tumour more frequent in dogs, counting for about 50% of the cases [1,3]. In human medicine, viral infection with hepatitis virus B or C, cirrhosis for alcohol consumption, and other condition (i.e. obesity) are reported to be the more frequent risk factors for HCC [4–6]. On the contrary, in veterinary medicine the viral aetiology has never been confirmed, and cirrhosis are rare.

HCC has more frequently (53-83%) a massive presentation, and in most cases is located into the left, medial or caudate (papillary process) liver lobe. Less frequently has a nodular (16-25%) or diffused (up to 19%) presentation. Regional lymph nodes and distal metastases (usually lungs and peritoneum) are less frequent for massive form (0-37%) than for nodular and diffused form (93-100%) [1,3].

Biliary system tumours

Biliary system tumours include biliary duct adenoma (BDA) and biliary duct carcinoma (BDC).

BDA is the second primary hepatic malignant tumour for incidence in dogs (22-41%). Usually is not associated to specific symptoms, nevertheless, when compression of the surrounding parenchyma occurs, some secondary effects could be detectable.

BDC is the second primary malignant tumours in dogs, with an incidence of 22-41%. In human being, a correlation between BDC and cholelithiasis, cholangitis and trematodes infestation are reported. The same risk factors have not been demonstrated in veterinary medicine [1,3].

In most cases (37-46%) BDC is characterised by massive morphology, followed by nodular (up to 54%) or diffused (17-54%) morphology. Furthermore, BDC is more frequently an intrahepatic tumour with aggressive biological behaviour. Indeed, the 88% of the patients present regional lymph nodes or lung metastases at the diagnosis, and carcinomatosis is present in 67-80% of the cases [1,3].

Neuroendocrine tumours

Primary neuroendocrine tumours (or carcinoids) are very rare in dogs. These tumours derive from neuroectoderma and the patient is typically a young patient if compared to HCC. Furthermore, primary neuroendocrine tumours are characterized by aggressive biological behaviour: they have diffused (67%) or nodular (33%) morphology, involving more than one hepatic lobe, and with local and distal metastases [1,3].

Mesenchymal tumours

Primary mesenchymal benign (haemangioma) and malignant (hemangiosarcoma, leiomyosarcoma, rhabdomyosarcoma, liposarcoma, osteosarcoma, malignant mesenchimoma) are rare in dogs. On the other hand, hemangiosarcoma metastases are frequent. Primary hepatic sarcomas are more frequently (36%) massive tumours or nodular (64%), with splenic and lung metastases reported in 86-100% of the cases [1].

Diagnostic imaging

Ultrasound

Ultrasonographic examination (US) is the first diagnostic step when a hepatic (or abdominal in general) pathology is suspected. No specific ultrasonographic features reported to be useful in the distinction among different hepatic pathologies leading to a low specificity of this diagnostic imaging technique. Therefore, a pathological investigation of the liver is nowadays paramount to reach a final diagnosis [2]. A possible criterion to distinguish benign from malignant lesions is the presence of abdominal fluid and dimensions of the lesion: indeed, nodules bigger than 3 cm are reported to have a high possibility of being malignant [2,7]. Similarly, hypoechoic nodules with hyperechoic core (the so call target-lesions) are more frequently metastatic nodules. Nevertheless, the same echotexture is reported also for nodular hyperplasia [8]. During US, is not possible to distinguish between HCA and HCC since both could be characterized by big, iso- or hyper-echoic, often cavitate lesions [8]. Hepatic and biliary cysts are easy to distinguish using US: they are characterized by anechoic, and well defined lesions, with distal acoustic enhancement, with normal surrounding parenchyma [2].

A synthesis of principal differential diagnosis for focal liver lesions, along with their US characteristics, is reported in Table 1.

Table 1 Differential diagnosis for more common focal liver lesions, along with their common US characteristics. Table from Penninck D., d'Anjou MA, Atlas of Small Animal Ultrasonography, second edition.

Anechoic lesions	Hypoechoic lesions	Hyperechoic lesions	Lesions with mixed echogenicity
Cyst	Nodular hyperplasia	Nodular hyperplasia	Nodular hyperplasia
Cystic tumours	Metastasis	Primary hepatic neoplasia	Primary hepatic neoplasia
Necrosis	Lymphoma	Metastasis	Metastasis
Abscess	Primary hepatic neoplasia	Mineralization, cholelithiasis	Abscess
Haematoma	Abscess	Abscess	Haematoma
	Necrosis	Myelolipoma	
	Haematoma	Granuloma	
	Complex cyst	Gas	

The contrast-enhanced ultrasonography (CEUS) diagnostic technique has increased the US accuracy. In human medicine, CEUS has become part of the procedures suggested by the guide lines for focal liver lesions, and for other organs lesions. Furthermore, the typical characteristics of malignant and benign FLLs are reported in the human literature [9,10]. The CEUS sensitivity for FLLs evaluation is reported to be comparable to computed tomography (TC) and magnetic resonance (MRI) [11]. For example, HCC is clearly identifiable during CEUS examination, being characterized by early wash-in hyperenhancement, late wash-out of the lesions, with non-enhancing areas inside the lesions representing haemorrhagic or necrotic areas [9].

During the last decades some papers regarding the application of CEUS in veterinary medicine have been published [12–15]. Nevertheless, even if some significant echocontrastographic features have been reported to be useful to distinguish benign and malignant FLLs (i.e. benign lesions are described as more frequently isoenhancing during the entire CEUS examination [13]), is not possible the distinction for HCC and BDC. Their CEUS features are, indeed, superimposable since they can exhibit all the possible enhancement [14].

Computed tomography

A specialty examination of FLLs through TC examination is often required for better investigate organs and lesions in a tridimensional way, for evaluation of tissues vascularization and lymph nodes reactivity, as well as local and distal metastases.

Nodular hyperplasia is reported to be more frequently characterised by hypoattenuating or hyperattenuating lesions during arterial phase, isodense to the radiologically normal liver parenchyma during pre-contrast, portal and delayed phase [16]. But HCAs show similar features. BDAs are usually characterised by the presence of cystic areas inside the lesions and by low enhancement, with peripheral distribution of the contrast medium [16].

HCC has typically cavitate appearance due to the presence of even big cystic areas that appear as hypoattenuating areas [16]. Nevertheless, also BDC is reported to show heterogeneous distribution of the contrast medium, with the possible presence of big non-enhancing areas [16]. Metastatic lesions can be both hypo- and hyper-attenuating, depending on their blood supply[16].

In human medicine, the tomographic features of liver masses are well known and described in the literature [11,17,18]. On the contrary, in veterinary medicine the tomographic features of these lesions are only seldom reported [19–25] and often with conflicting results. Indeed, in the studies of Griebie et al 2017 and of Burti et al 2021 the enhancement during delayed phase and maximum dimensions of the lesions are reported to be statistically significant in the distinction between benign and malignant FLLs. On the contrary, Stehlìk et al 2020 demonstrated that none of the tomographic features can be useful for the distinction of any type of lesion.

In Figures 1-9 some examples of focal liver lesions as they appear during US and CT examination are reported.

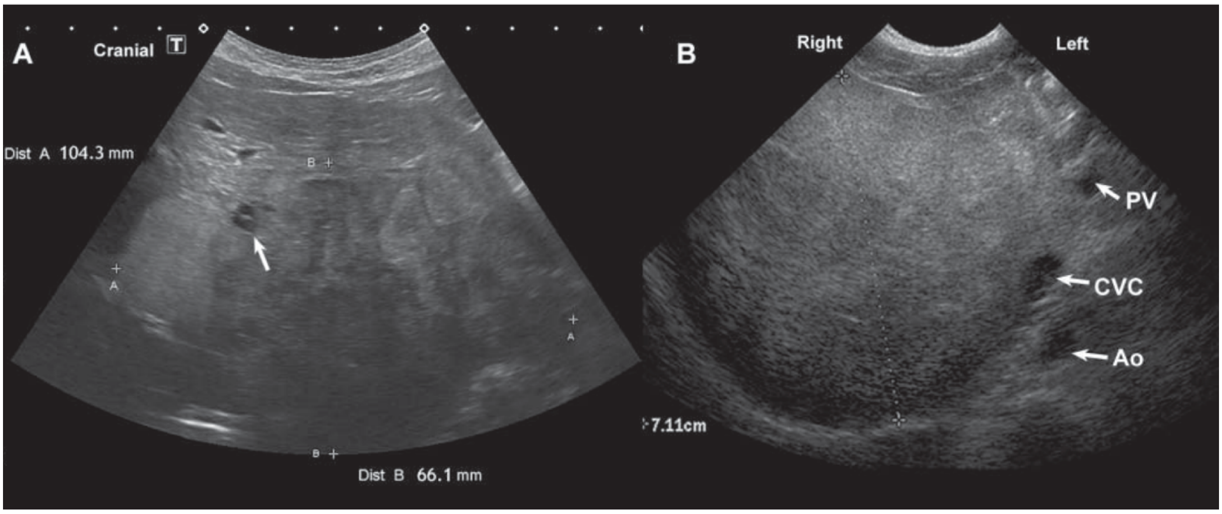


Figure 1 US images of two cases of hepatoma. A. mass characterised by irregular and not well-defined margins, moderately echogenic and heterogeneous, with small cavitate areas. B. Hyperechoic heterogeneous mass. PV, portal vein; CVC, caudal vena cava; Ao, aorta. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

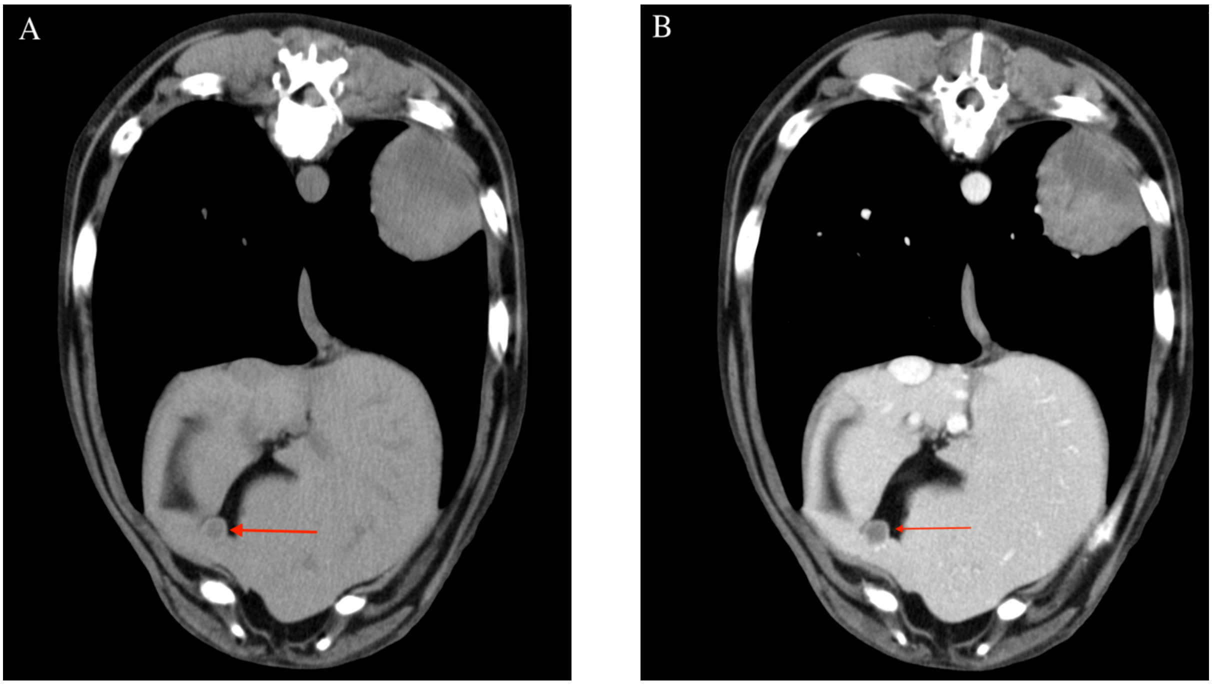


Figure 2 CT images of a case of hepatoma. A. Pre-contrast phase, the lesion shows hypoattenuation, well-defined margins and hyperattenuating capsule. B. During delayed phase the lesion is hypoattenuating.

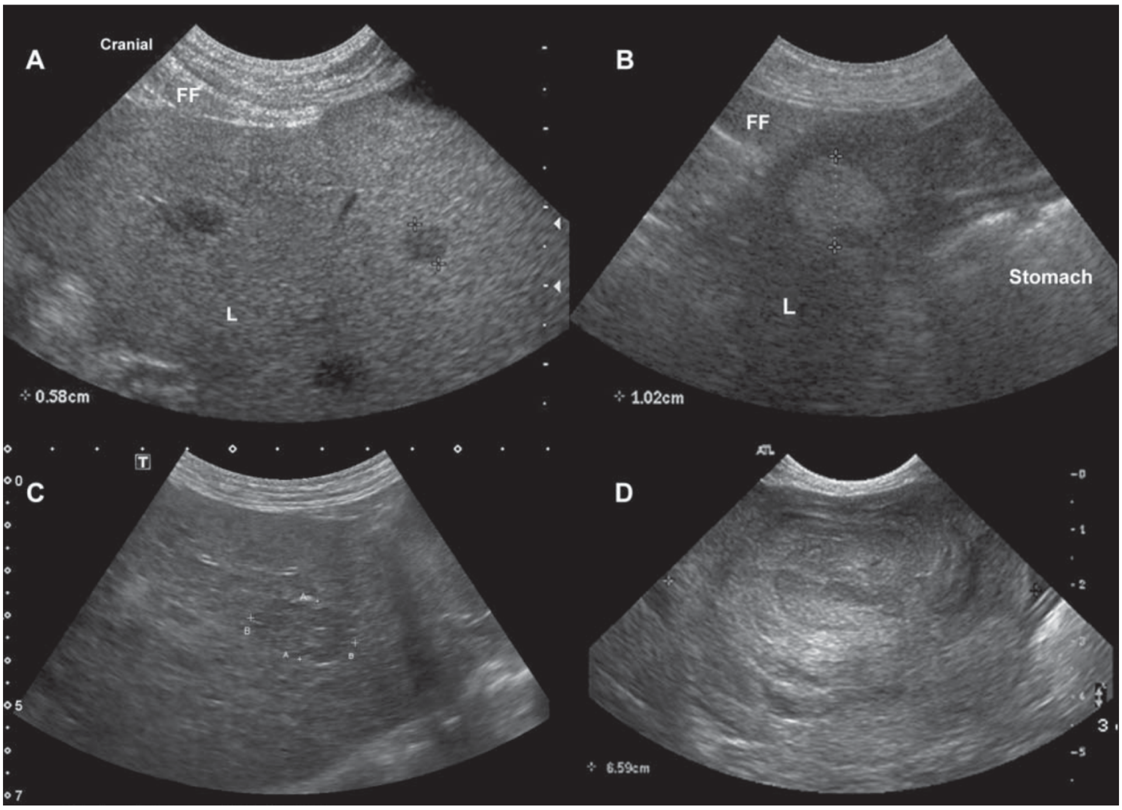


Figure 3 US images of 4 cases of nodular hyperplasia. A. Small hypoechoic lesions, with ill-defined margins. B. Hyperechoic and well defined lesion. C. Heterogeneous, hypoechoic lesion, with ill-defined margins. D. Heterogeneous mass. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

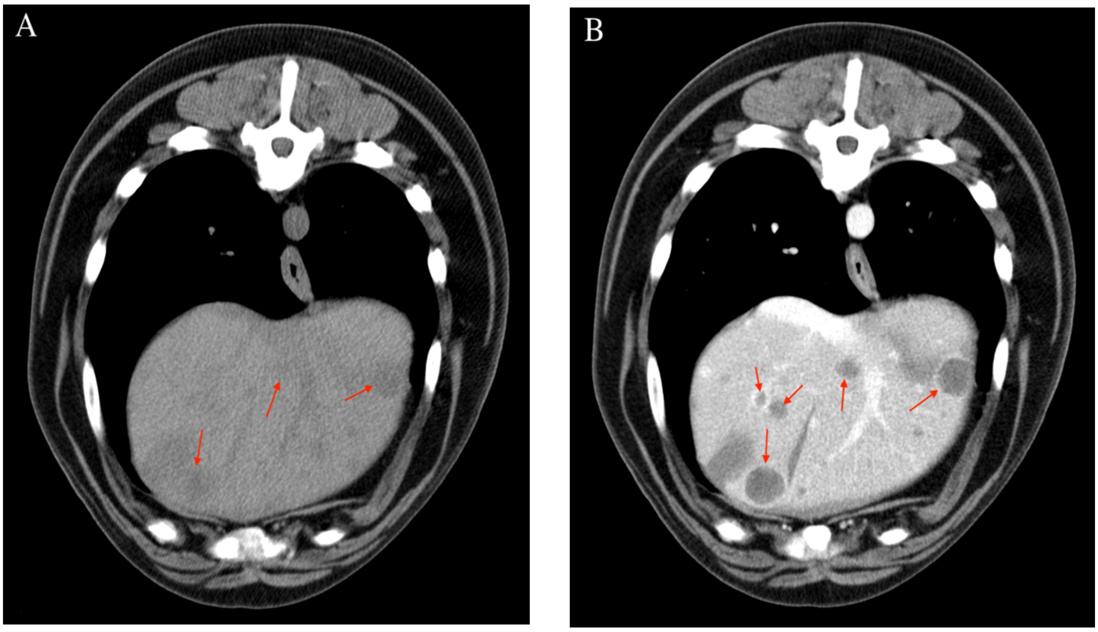


Figure 4 CT images of nodular hyperplasia. A. Pre-contrast phase, nodular hyperplasia appear as slightly hypoattenuating lesions, with ill-defined margins. B. Delayed phase, nodular hyperplasia appear as hyoattenuating lesions, with well-defined margins.

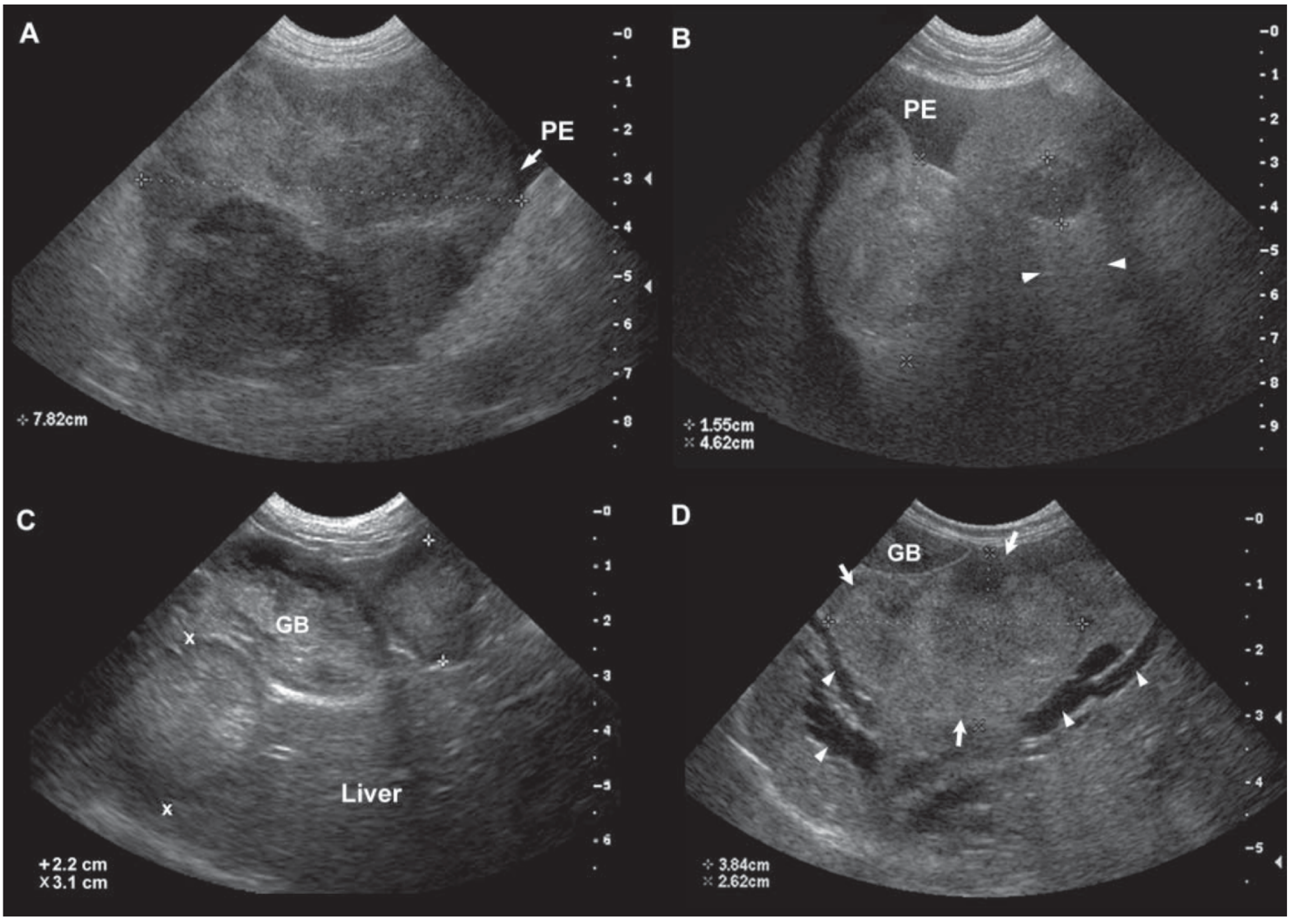


Figure 5 US images of four examples of primary malignant liver tumours. A. Hepatocellular carcinoma, heterogeneous massive lesion, with irregular and ill-defined margins. B. Hepatic adenocarcinoma, visible as a rounded hyperechoic lesion, with presence of an other hypoechoic nodule characterised by distal acoustic enhancement (PE, peritoneal effusion). C. Neuroendocrine carcinoma, characterised by hyperechoic nodules with hypoechoic margins (GB, gallbladder). D. Biliary duct carcinoma (in a cat) characterised by a lobulated, hyperechoic and heterogeneous mass. The tubular anechoic structures (arrow head) are dilated biliary ducts (GB, gallbladder). Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

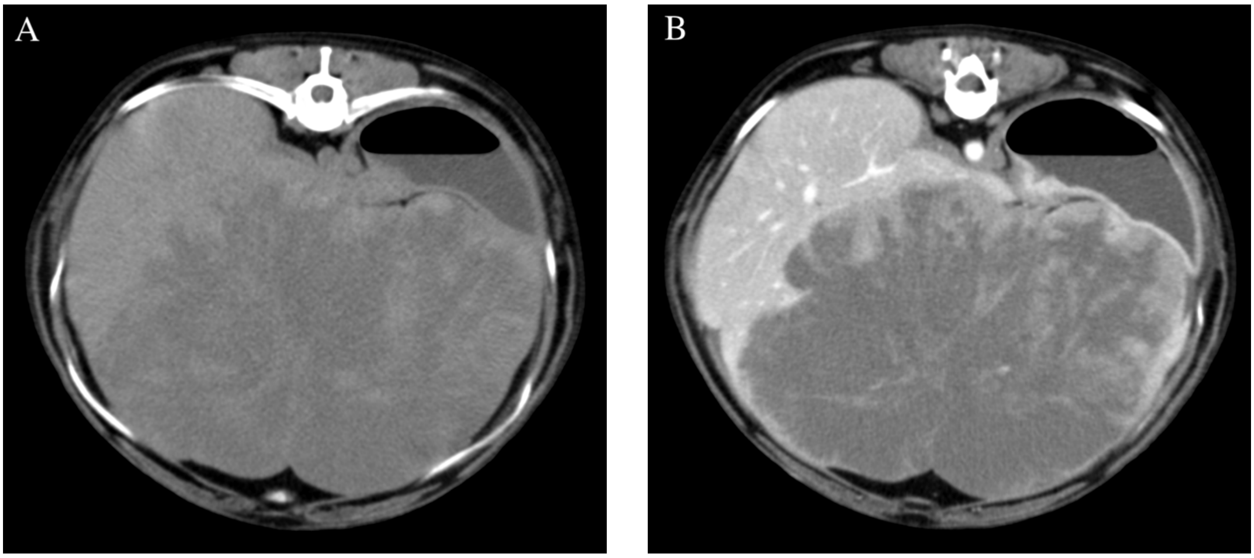


Figure 6 CT images of a case of massive hepatocarcinoma that appear as: A. hypoattenuating lesion, with ill-defined margins, during pre-contrast phase; B. hypoattenuating and heterogeneous lesion, with ill-defined and irregular margins.

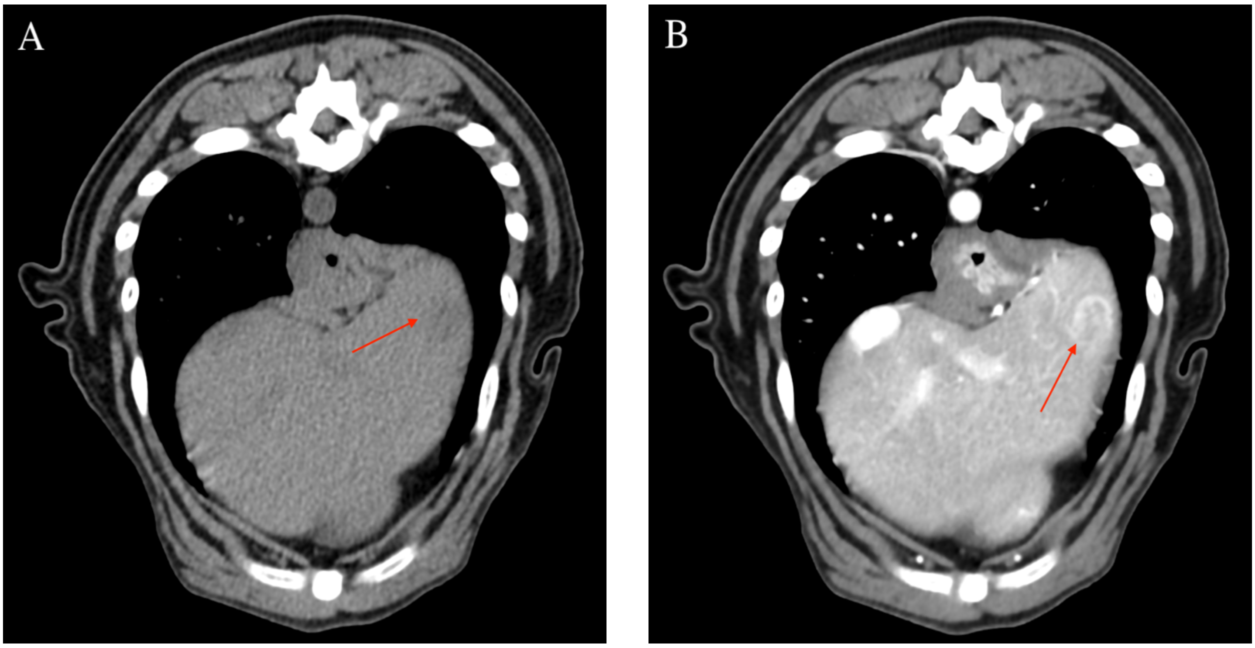


Figure 7 CT images of a case of neuroendocrine neoplasia of the liver that appears as: A. slightly hypoattenuating lesion during pre-contrast phase, with ill-defined margins; B. hyperattenuating and heterogeneous lesion during delayed phase, with irregular and well defined margins.

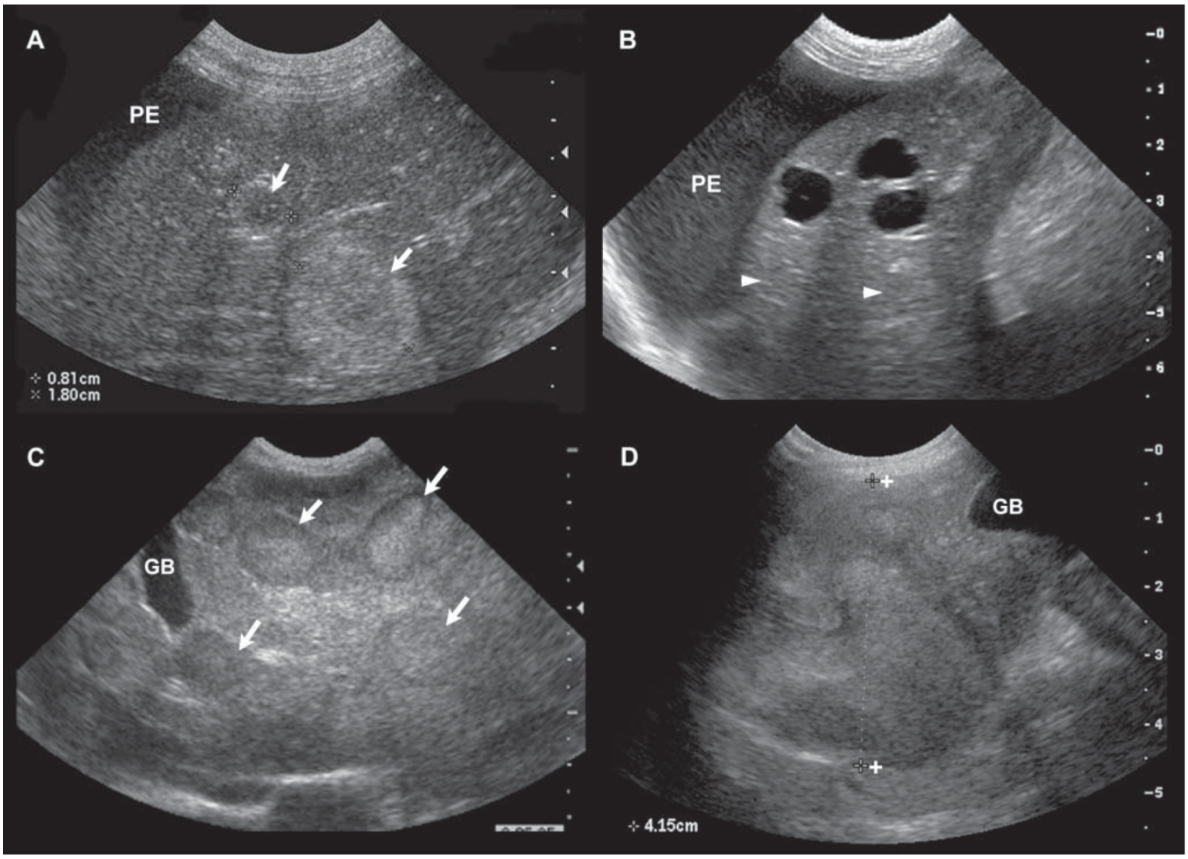


Figure 8 US images of four cases of liver metastases from: A. pancreatic adenocarcinoma, two nodules are visible, the smaller is hypoechoic, with thin hyperechoic margins, the biggest is hyperechoic (PE, peritoneal effusion); B. splenic haemangiosarcoma, visible as cavitate lesion, with anechoic areas characterised by acoustic enhancement; C. thyroid carcinoma, characterised by target lesions; D. pancreatic adenocarinoma, visible as heterogeneous hyperechoic mass. GB, gallbladder. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

Figure 9 CT images of two cases of metastases. A. pre-contrast phase of neuroendocrine metastastatic tumour, the lesions appears hypoattenuating, heterogeneous and with irregular margins; B. the same lesion during dalayed phase appears hypoenhancing; C. pre-contrast phase of metastatic melanoma that appear slightly hypoattenuating; D. the same lesion during delayed phase appear as hypoenhancing lesion, with irregular margins.

FOCAL SPLENIC LESIONS IN DOGS

Focal splenic lesions (FSLs) are a common finding in elderly dogs, and are mostly benign, even if complication such as hemoabdomen due to incidental ruptures might occur [3,26,27]. Neoplastic secondarisms to the spleen are uncommon, counting for about 1-6% of all the splenic neoplasia. Usually metastatic process that involved the spleen derived from sarcomas, carcinomas or neuroendocrine tumours [3,26,27].

Vascular lesions

Splenic lesions originating from the vascular components of the spleen are usually malignant, but might also be benign, and no specific features morphological and clinical features can be used in the distinction between benign and malignant lesions. As a consequence, the distinction among different types of vascular lesions can be really challenging [3]. Hematomas and haemangiomas are the most common benign vascular tumours. Hematoma is a rupture of a portion of the splenic parenchyma, with consequent intraparenchymal haemorrhage. The cause is often idiopathic or benign as amyloidosis, infections, nodular hyperplasia, or trauma. Nevertheless, a hematoma could occur also as a consequence of the rupture of a sarcoma, especially if haemangiosarcoma.

Haemangioma has a low prevalence in dogs (1-5%)[3], but the morphological features are very similar to those of splenic haemangiosarcoma.

Haemangiosarcoma is the more commonly diagnosed splenic malignancy in dogs between 6 and 17 years old, has a high metastatic rate, and an unfavourable prognosis. The life expectancy after the diagnosis is less than one year; if haemoabdomen is present at the time of the diagnosis the life expectancy is reduced to four months [3].

Mesenchymal tumours

Splenic mesenchymal tumours is a heterogeneous group of neoplasia that overall comprehend the 25-50% of all the splenic neoplasia. This group includes both benign (nodular hyperplasia, fibroma, lipoma, mielolipoma, and leiomyosarcoma) and malignant (fibroma, leiomyosarcoma, liposarcoma, myosarcoma, rhabdomyosarcoma, chondrosarcoma, and osteosarcoma) neoplasms.

Nodular hyperplasia (NH) is a common benign splenic proliferation in dogs that might originate from different splenic cellular types (haemopoietic, lymphoid or complex component)[3,26]. NH does not cause clinical symptoms and is an occasional finding in US, TC or necroscopic examinations.

Malignant mesenchymal tumours have an aggressive biological behaviour, with a high metastatic rate (70%) and an unfavourable prognosis. The survival rate ranges between 1 month (if metastases are present) and 9 months (if a single mass is present). The liver is the most common site of metastasis (85% of the cases), but also peritoneum, mesentery and lung might be affected[3].

Diagnostic imaging

Ultrasound

Focal splenic lesions are a frequent finding in abdominal ultrasonographic examinations. Like for liver lesions no US features are reported to be useful in the distinction among different lesion histotypes, Table 2. For example, nodular hyperplasia I reported to be hypoechoic, but also hyperechoic. Target lesions are likely to be malignant lesions, but also nodular hyperplasia and extramedullary hemopoiesis can display this type of echotexture. Honey-comb echotexture of the spleen (diffused submillimetrical hypoechoic nodules) is suggestive of lymphoma. Nevertheless, is also reported for other malignancies (i.e. metastases of adenocarcinoma). Myelolipomas are more frequently localized on the mesenteric surface of the spleen, are hyperechoic with regular and well defined margins, distal acoustic shadow might sometimes be present [29].

The US features of big cavitary lesions with haemorrhagic/necrotic areas are often similar to those of big hematomas.

Table 2 Differential diagnosis for more common focal splenic lesion, along with their common US characteristics. Table from Penninck D., d'Anjou MA, Atlas of Small Animal Ultrasonography, second edition

Hyperechoic lesions	**Hypoechoic lesions**	**Lesions with mixed echogenicity**	**Anechoic lesions**
Myelolipoma	Nodular hyperplasia	Nodular hyperplasia	Cyst
Haematoma	Extramedullary haematopoiesis	Extramedullary haematopoiesis	Abscess
Nodular hyperplasia	Haematoma	Haematoma	Splenic pseudocyst
Extramedullary haematopoiesis	Abscess	Abscess	
Granuloma	Neoplasia	Neoplasia	
Abscess	**Multifocal hypoechoic nodules**	**Target lesions**	
Neoplasia	**(honey comb aspect)**	Metastasis	
	Lymphoma	Extramedullary haematopoiesis	
	Histiocitic sarcoma	Nodular hyperplasia	
	Extramedullary haematopoiesis		
	Cuneifon hypoechoic areas		
	Infarction		

The application of CEUS for the evaluation of FSLs is still scarcely evaluated in veterinary medicine. In a study by Ohlerth et al 2008 all the enhancement pattern for the benign FSLs during CEUS examination are reported. Vice versa, malignant FSLs are described as markedly hyperenhancement [30]. On the contrary, Nakamura et al 2010 described malignant FSLs as being more commonly hypoenhancing [31].

Computed tomography

Computed tomographic features of FSLs are poorly reported in the veterinary medicinal literature, and the results are inconsistent. Benign FSLs are reported to mainly have diffused and uniform distribution of the contrast medium, with an attenuation pattern similar to those of the radiologically normal splenic parenchyma [16]. Nevertheless, nodular hyperplasia is reported also to be hyperenhancing in post-contrast phase. Extramedullary hemopoiesis is reported as markedly hyperenhancing in all the post-contrast phases. Haematoma is reported to have an heterogeneous appearance [32–35].

Among malignant lesions, haemangiosarcoma is mainly characterized by having a complex mass appearance, with hypo- or non-attenuating pattern. In some occasions hemangiosarcoma might have the same enhancing pattern of the surrounding splenic parenchyma [16].

In a study of Fife et al 2004 a cut-off of 55 Hounsfield Unit (HU) is reported to be useful in the distinction between malignant and benign FSLs. Nevertheless, other studies found different cut-off values.

In Figures 10-17 some examples of focal splenic lesions as they appear during US and CT examination are reported.

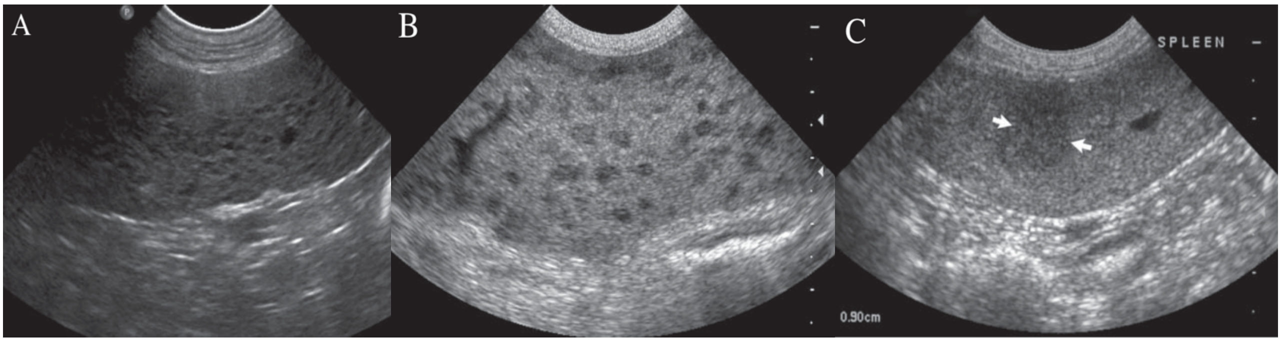


Figure 10 US images of 3 different examples of lymphoma characterised by: A. small hypoechoic diffused nodules; B. hypoechoic nodules diffused, the lesion is slightly bigger compared to previous example; C. a single hypoechoic nodule is visible. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

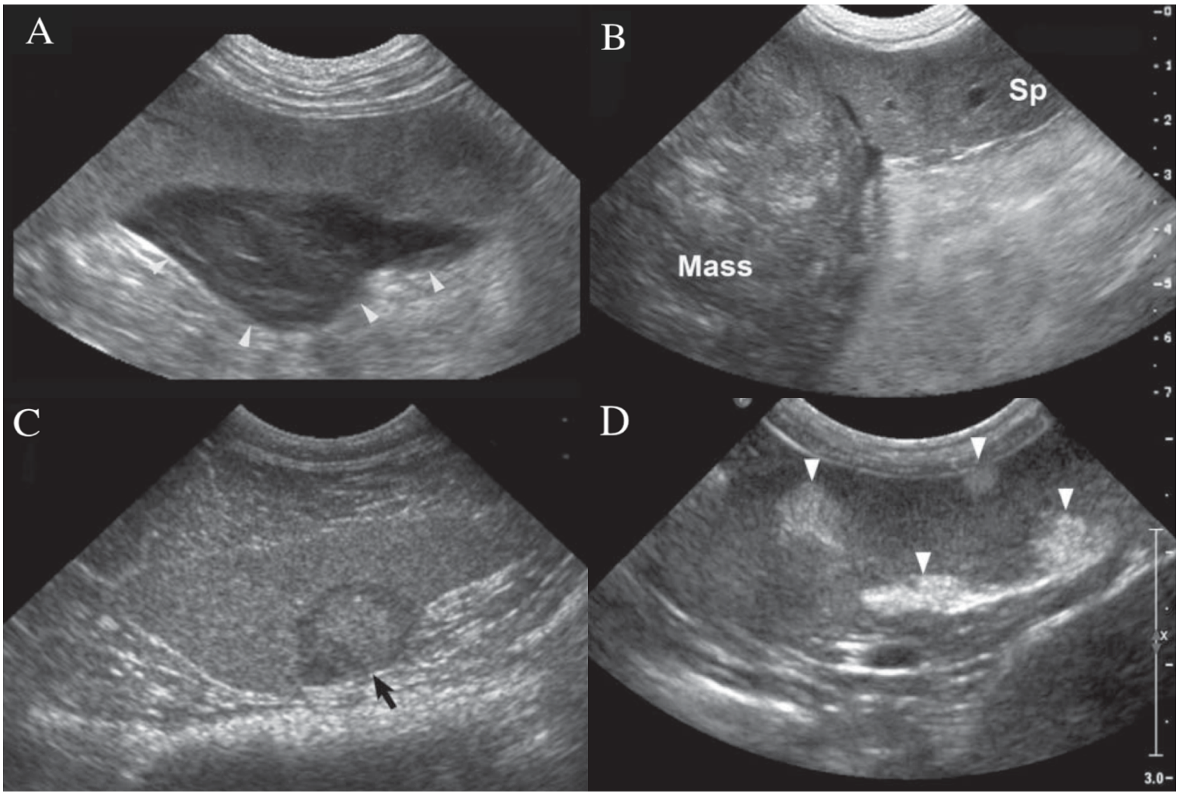


Figure 11 US images of four different examples of benign lesions: A. splenic hematoma, that shows hypoechoic appearance; B. nodular hyperplasia, characterized by a large mass with irregular echogenicity, and by regular and well-defined margins; C. extramedullary haematopoiesis, that appears as a hypoechoic and heterogeneous lesion, with regular and well-defined margins; D. myelolipomas (arrow heads), strongly hyperechoic and acoustic shadow in the left nodule. Figures from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

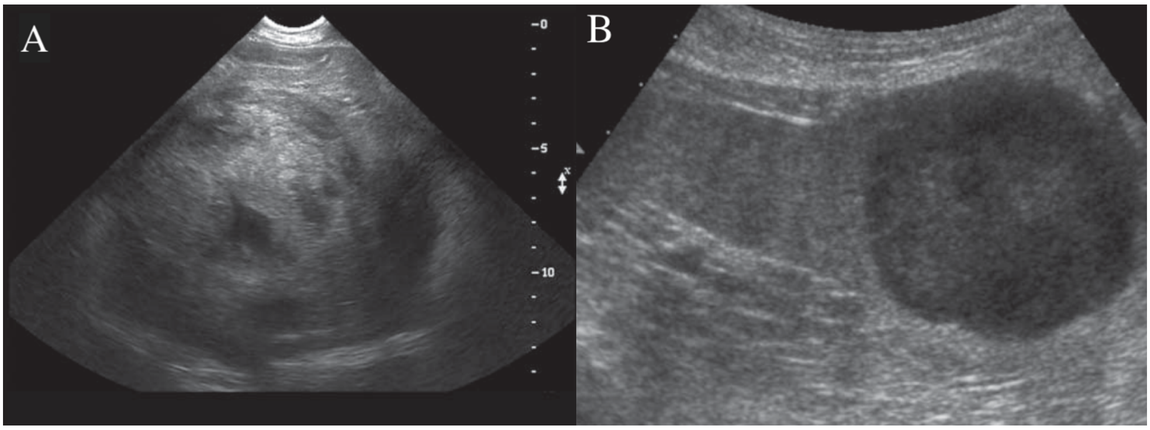


Figure 12 US images of 2 examples of benign lesions: A. nodular hyperplasia characterized by mixed echogenicity, hypoechoic areas, and well-defined margins; B. pyogranulomatous splenitis, characterized by a single hypoechoic mass. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

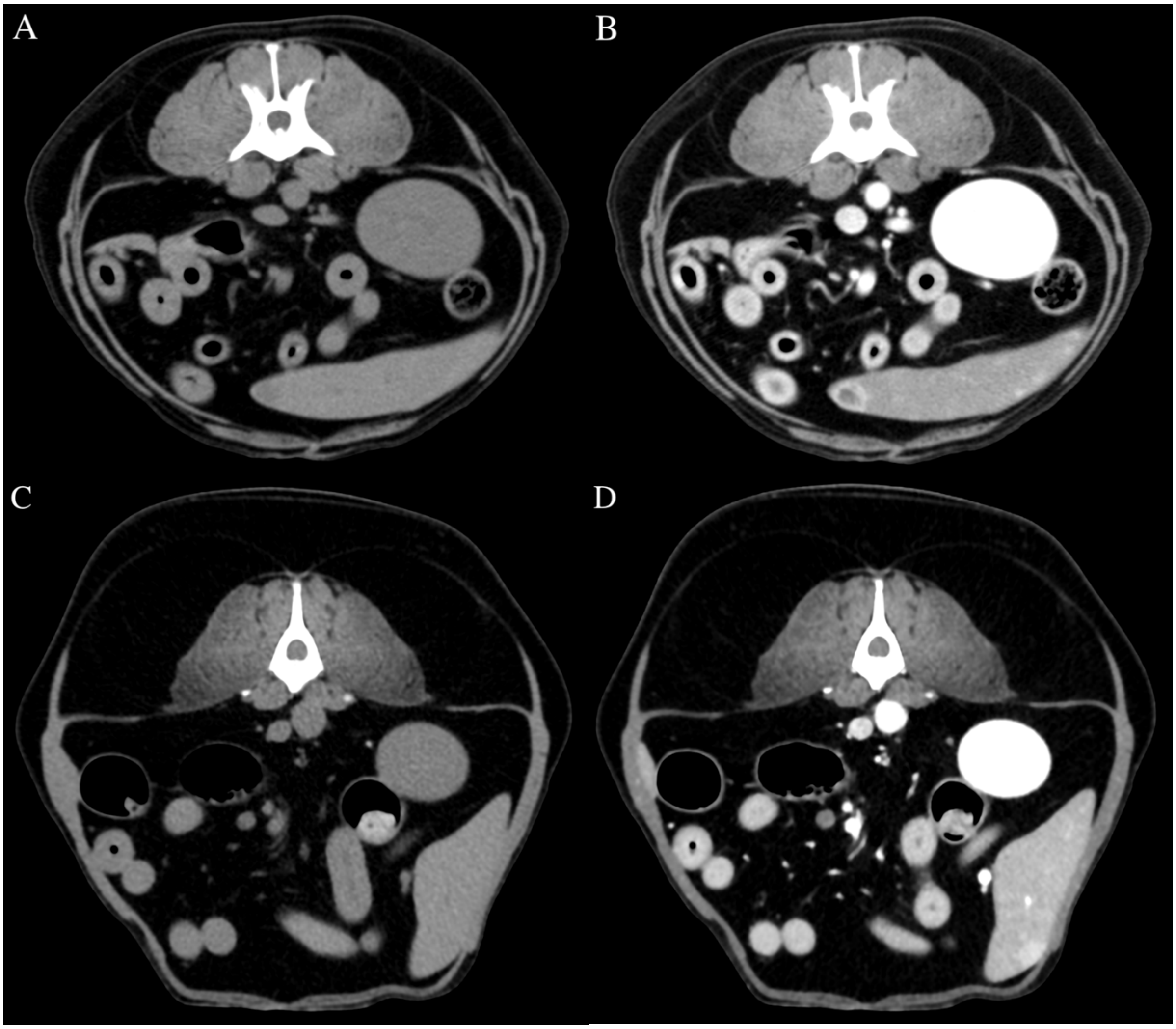


Figure 13 CT images of 2 different type of benign lesions. In the left column (figures A, C) pre-contrast images are reported; in the right column (figures B, D) delayed phase images are reported. A. and B. nodular hyperplasia, isoattenuating during pre-contrast phase, hypoattenuating lesion with hyperattenuating and well-defined margins during delayed phase; C. and D. myelolipoma, isoattenuating during pre-contrast phase, hyperattenuating during delayed-phase.

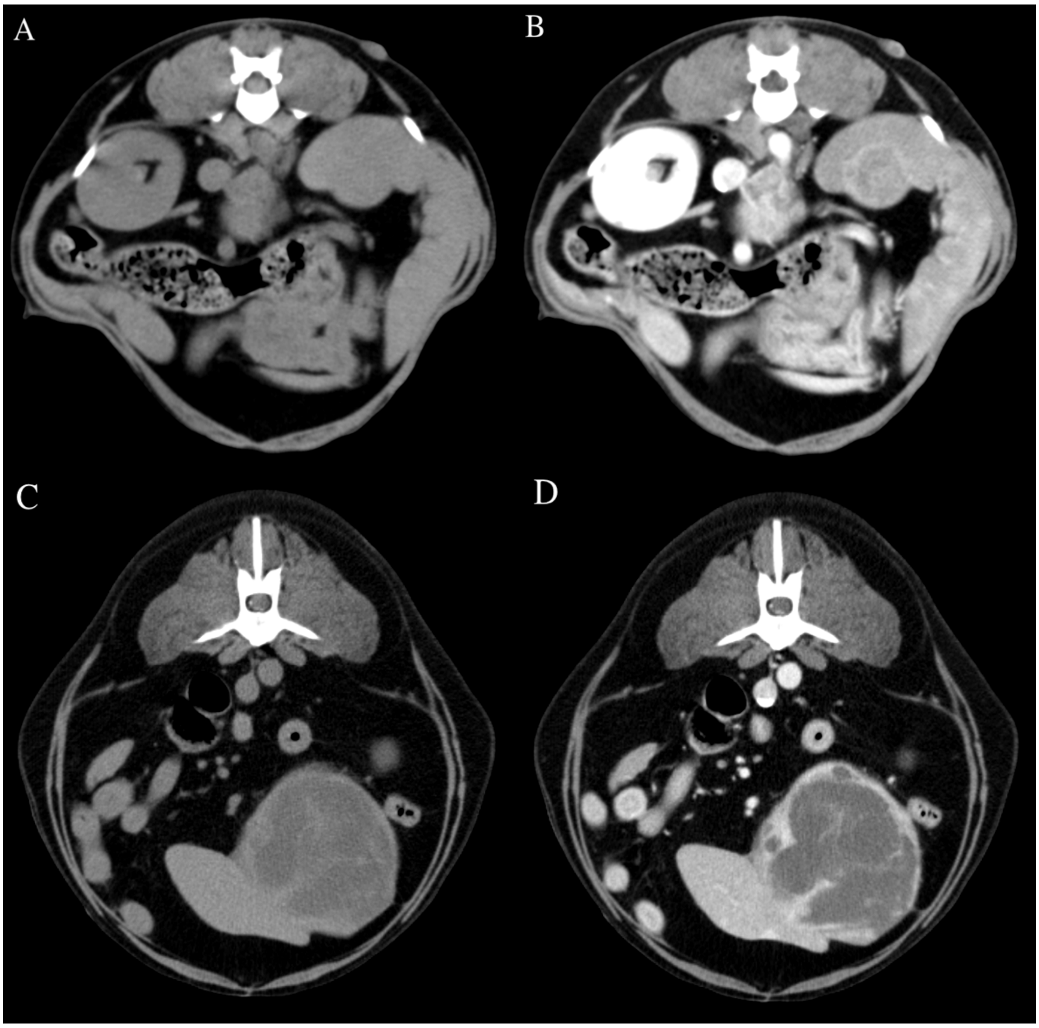


Figure 14 CT images of 2 different type of benign lesions. In the left column (figures A, C) pre-contrast images are reported; in the right column (figures B, D) delayed phase images are reported. A. and B. extramedullary haematopoiesis, slightly hypoattenuating in pre-contrast phase, hyperattenuating and heterogeneous, with irregular and well-defined margins during delayed-phase; C. and D. haematoma, that appear as a big hypoattenuating lesions during pre-contrast phase, hypoattenuating and with hyperattenuating septa during delayed phase.

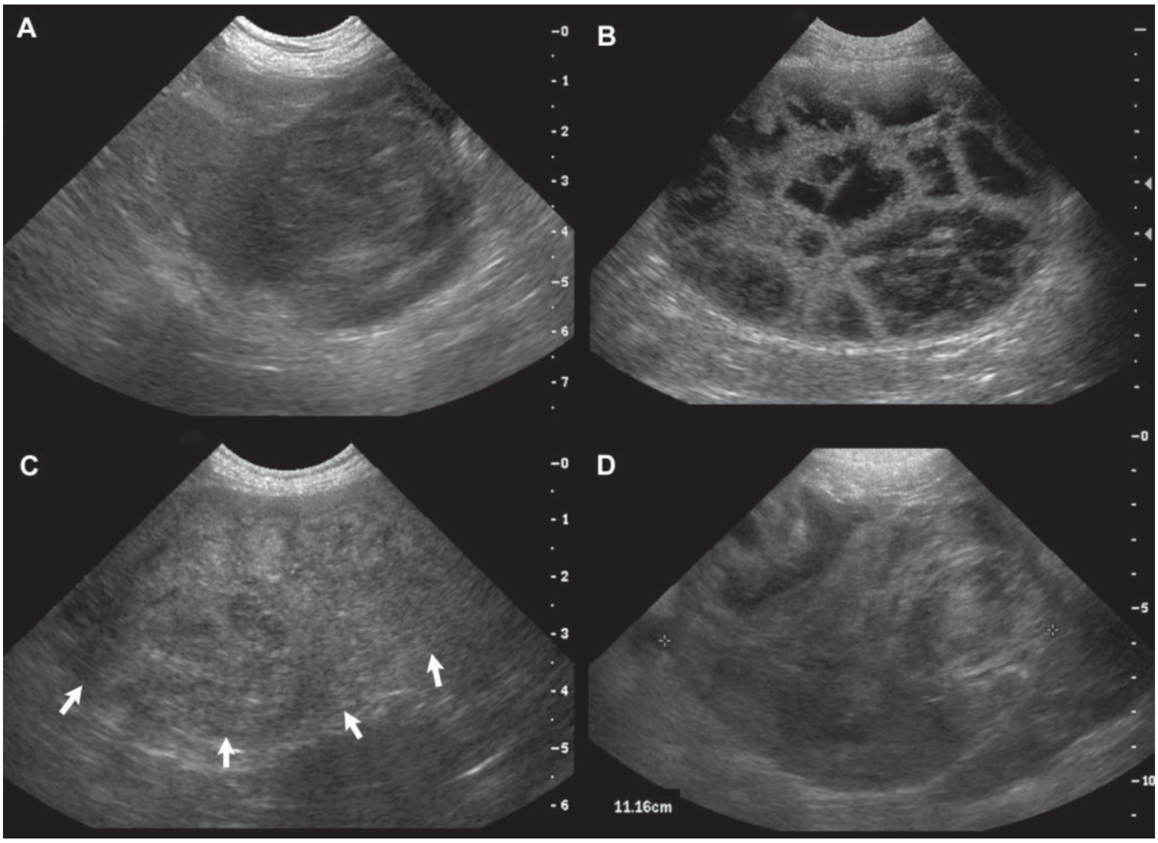


Figure 15 US images of 4 cases of haemangiosarcoma: A. mass characterized by mixed echogenicity; B. big mass with well-defined margins, heterogeneous aspect, and hypo/anechoic cavitations; C. heterogeneous mass with not defined margins; D. mass with mixed echogenicity. Figure from Atlas of small animal ultrasonography, Penninck D, d'Anjou MA, Wiley Blackwell, 2015

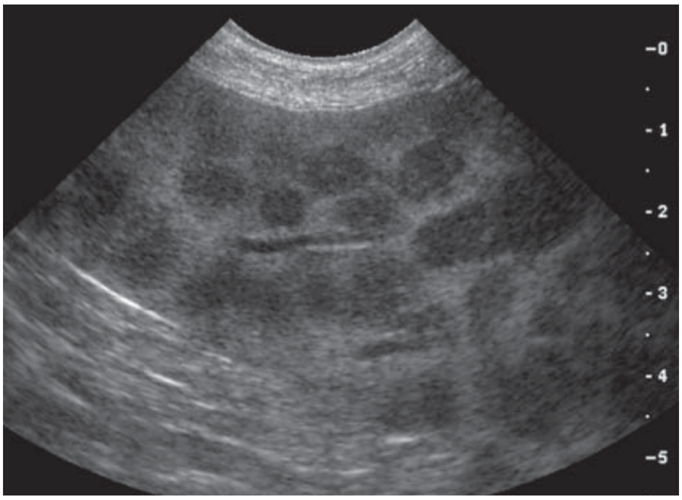


Figure 16 US image of a case of metastases from anal sac adenocarcinoma. Numerous hypoechoic diffused nodules are visible. Figure from Atlas of small animal ultrasonography Penninck D, d'Anjou MA, Wiley Blackwell, 2015

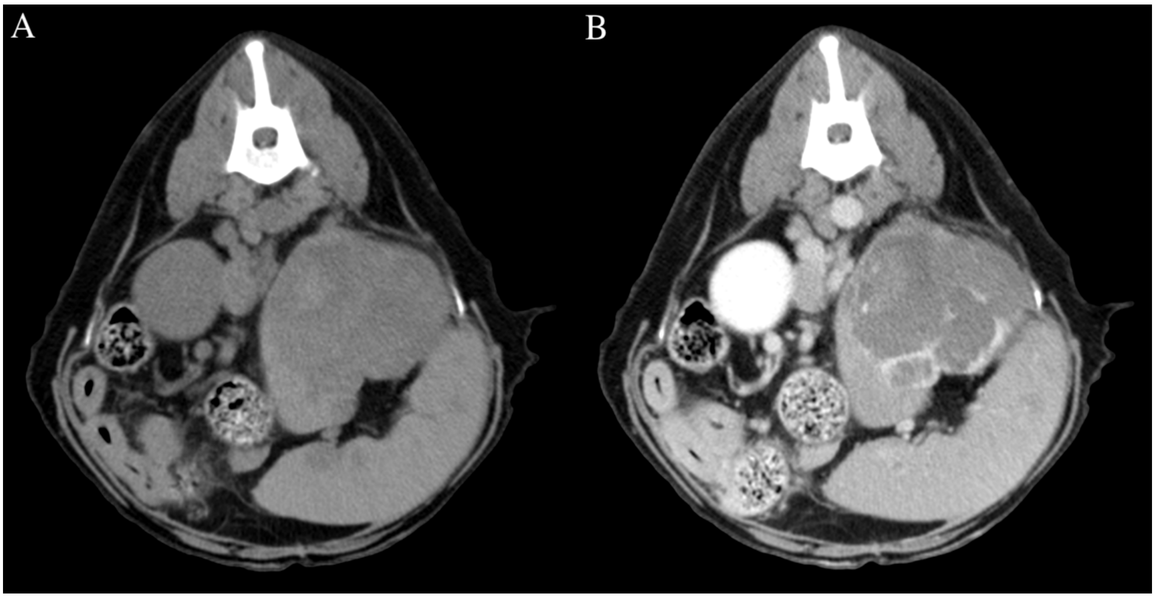


Figure 17 CT images of primary splenic haemangiosarcoma, showing hypoattenuation in both pre-contrast (A) and delayed phases (B). Hyperattenuating septa are visible in delayed phase.

DEEP-LEARNING AND ITS APPLICATIONS

Deep learning

In the last decades, artificial intelligence has seen considerable improvements, visible also in various aspect of every-days life. Today, deep learning-based technologies are able to solve tasks like objects classification, speech recognition, language detection, automatic driving of drones and cars, image processing, interaction with home entertainment system, and much more. The first algorithms working with simple neural network date back to the ’40s and ’50s, when Warren McCulloch and Walter Pitts first and the psychologist Frank Rosenblatt after, pioneered this field [36]. During the following decades, more and more studies were conducted, but the advances in the development of this technology were hindered by the limited computing power available at that time.

Only in 2006 the application of deep-learning saw a leap forward with the availability of bigger and bigger dataset, and with the introduction of high-performance graphic processing units (GPU) that can allowed faster training of complex neural network.

Nowadays, there is still come confusion regarding the differences existing between artificial intelligence, machine learning, and deep-learning [37]. In Figure 18 a chart that demonstrates the relationship among them is reported.

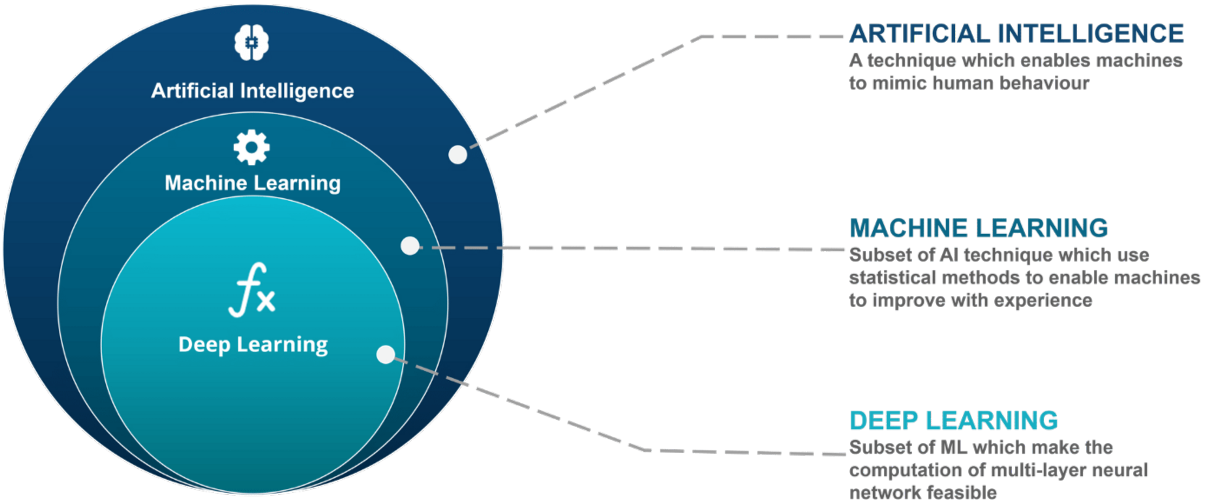


Figure 18 A chart that demonstrates the relationship between AI, ML, and DL.

Artificial intelligence

Artificial intelligence (AI) is a field of the science were technologies capable to solve tasks that typically required the presence of the human being are developed. Choi et al 2020 gave a very explanatory example of this, citing a hypothetic AI algorithm programmed for the room temperature control. The person who schedules this control-system already knows the temperature ranges that subjectively guarantee the optimal room comfort, so he decides a threshold: when the room temperature rises the control-system works to cool; vice versa, when the room temperature has lowered the control-system works to warm. This kind of AI is not able to recognize people or objects.

Machine learning

Machine-learning (ML) is a subfield of AI. The term ML refers to algorithms capable of solving complex tasks without being explicitly programmed [38]. Furthermore, these algorithms can be used to make decisions based unseen data. Given an initial target, ML systems starts from specific features to solve the task. The features are usually manually assigned by an operator [38].

There are basically four classes of ML [37]:

- Supervised learning (i.e. decision-tree, logistic regression, linear regression), where a labelled dataset is used to train an algorithm to solve a specific task. The model is fed with the input data and then, though through subsequent adjustments, the model finds a proper way to classify the original dataset;
- Unsupervised learning (i.e. clustering, data analysis, image recognition), works with unlabelled dataset. In this case, the algorithms are able to find hidden patterns in the original

dataset, that the model uses to find similarities and differences, without human intervention. These patterns could not exist in the original dataset and are not expressed in the output: they are unknown. The interpretation of the outputs of these algorithms might be challenging.

- Semisupervised learning, where the original dataset includes both labelled and unlabelled images. This typically happens with dataset of medical images: entering the labels is a time-consuming process, and, as a consequence, the radiologist assigns the labels only to a part of the images. These images are used to train the model, that will be used to classify the remaining unlabelled images;
- Reinforcement learning: these are algorithm that learn not only from data, but also from subsequent trials and errors. This type of ML is still unused in medical field.

Convolutional neural networks and deep-learning

The neural network is a network of neurons communicating with each mimiking human visive and auditory cortex. Indeed, a neural network is composed of cellular bodies (the so call nodes) that are organized in layers. Data are inputted, analysed by the nodes of the first layer, and then transmitted to the following layer and so on. In every layer, features that will be useful for the representation of the original image given as initial input are extracted [37,39].

Deep learning (DL) is a subfield of ML consisting in neural networks with three or more layers: the first receives the input data, the last gives the output and, in the middle, there are a variable number of hidden layers. Thereafter, the predicted class of the imput data is compared to the actual class. The difference between the predicted and the actual class is computed and the weights of each layer are iteratively adjusted to minimize the difference until no further improvement is obtained.

There are different types of neural networks ((*artificial neural network* (ANN), *recurrent neural network* (RNN), *generative adversial neural networks* (GANN) among the more used). Nevertheless, in the radiology field, convolutional neural networks (CNN) are the more applied.

CNNs became famous in the ‘90s with a milestone study by LeCun et al 1998. Let’s assume that we need to recognize an image. In the CNN, the nodes of one layer identify some areas, of different dimensions, inside the image, the so call patches. Every patch is partially superimposed to other nearby patches, so that the proper spatial context is preserved: patches that are located in nearby positions in the original image share more information that patches located in distant positions. The CNN identifies and focuses the features of every patch and create some filters, the so call convolutional filters. The convolutional filter passes over all the image in a succession of mathematical multiplications called convolution. The output of every convolution is a feature map, a matrix containing the numerical data corresponding to where the convolutional filter has found or not found a specific feature. For example, if we want to find every horizontal line inside an image, the CNN will be trained for this task and will create a feature map with different values referred to where there is or there is not a horizontal line.

All this complicated process happens in every layer. At the end, all the feature maps will be compressed and sent to a classical ANN for the image classification, based on the extracted features [37,39]. It is easy to understand that, the deeper the network (that is, the more layer the network is composed of), the more numerous and detailed features would be extracted. In Figure 19 is reported a schematic representation of a basic CNN: the input coming from the tomographic image is given to the network for the training phase (the training will be composed of various cycles), during what the network itself extracts the features that allowed the image recognition; at the end, the network will give the optimal classification as output on the basis of the extracted features.

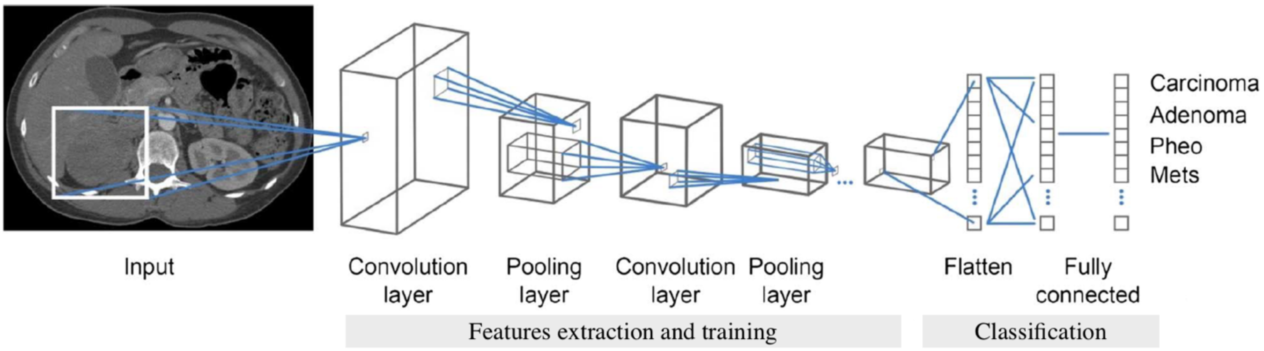


Figure 19 Schematic representation of a CNN. Figure from Moawad AW, Fuentes DT, ElBanan MG, Shalaby AS, Guccione J, Kamel S, Jensen CT, Elsayes KM. Artificial Intelligence in Diagnostic Radiology: Where Do We Stand, Challenges, and Opportunities. J Comput Assist Tomogr. 2022 Jan-Feb 01;46(1):78-90. doi: 10.1097/RCT.0000000000001247. PMID: 35027520.

The success of the DL compared to the traditional ML methods is basically related to 2 factors strictly linked each other: the deep and the compositionality of the network. The deeper the architecture, the smaller the training set can be, because starting from a smaller number of examples the network will be able to learn and to generalize on an unseen dataset. This is more important as more complex the task is and the smaller the original dataset is.

Today, there are different types of CNN based on different architectures, with various number of layers. The oldest is LeNet, born in 1998, that represents the skeleton for the following CNNs: AlexNet, VGGNet/OxfordNet, GoogleNet, DenseNet e ResNet.

Applications of deep-learning in Diagnostic imaging

The possibilities offered by the application of neural networks in the diagnostic imaging field, is mainly related to two factors:

1) The interpretation of diagnostic images is usually performed by a specialized radiologist but, in emergency situations, also less skilled veterinarians might need to interpret those images. Nevertheless, interpretative errors occur in both situation;
2) In case of liver or splenic focal lesions obtaining a final diagnosis is paramount for an adequate and prompt treatment for the patient. Nowadays, histological examination is the gold standard to achieve a definitive diagnosis, but, unfortunately, is not always possible for the patient to undergo a so invasive examination.

Errors in diagnostic imaging

Since its birth, diagnostic imaging assumed a central role in the clinical practice both in human and veterinary medicine. Nevertheless, even if experience, knowledge and technology has progessed during the last decades, the incidence of errors in diagnostic imaging is still high, with consequent delay in treatments and longer hospitalizations [40]. In veterinary medicine the incidence of these errors is still unknown. On the contrary, in human medicine several studies reported a frequency of diagnostic imaging errors ranging from 10 to 15%[41,42] for thoracic radiographic studies, and from 3 to 16% for abdominal tomographic studies [40,43].

The aetiology of such errors is still debated, but seems to be multifactorial: mechanical, physiological and psychological factors are simultaneously involved in a process still not deeply understand [44]. There are two fundamental errors in radiology: perception and interpretation errors [40].

The perceptive errors is when a lesion is missed during the initial evaluation of the images, followed by its identification in retrospective images evaluation [40,41]. In a study by Yun et al 2017 the causes of perception errors are summarized as follow: neglect the presence of other lesions once one abnormality is identified; presence of a bigger lesion near to a missed lesion; localization of the missed lesion in an anatomical blind region, that is, anatomical sites (more frequently liver, peritoneum, body wall, retroperitoneum and mesentery) that are not accurately observed by the radiologist maybe for cognitive reasons; use of an inappropriate window of visualization [40,45].

Interpretation errors is the incorrect diagnosis of an otherwise correctly identified lesion. These errors usually arise from a the inexperience of the radiologist (i.e. the classification of a normal tissue as abnormal or the incorrect interpretation of an abnormal presentation of a pathology in absence its typical features), but might also derive from an inappropriate collection of patient anamnesis [40].

In tomographic studies technical factors related to scan protocols, soma of the patient, contrast medium, possibilities to perform scans in different phases (arterial, venous, delayed) must also be taken in account[43].

Knowledge of the different error types is useful to develop adequate support systems for the radiologist. Some support systems are, for example, the use of structured reports, the reduction of multitasking and double evaluation of the medical images. Even if the constant use of these support systems is useful in the reduction of error rate, other, more sophisticated approaches are required to achieve optimal standards.

In the last decades, lot of attention was given to computed-aided detection (CAD) systems, where AI algorithms assist the clinician in the evaluation of radiographic, tomographic, and ultrasonographic images. In human medicine, several studies demonstrate the efficacy and utility of these methods and a plethora of possible applications have been advised [46,47].

Some CADs for diagnostic imaging have been also proposed in veterinary medicine, with encouraging results [48–53]. Among these, in the study by Burti et al 2020 four models based on four different CNNs for the evaluation of canine cardiomegaly in thoracic radiographies have been studied. All the models had a high diagnostic accuracy (area under the curve (AUC) > 0.9). Moreover, in the study by Banzato et al 2021 two CNNs have been applied to thoracic radiographs of the dog to detect some of the most commonly encountered thoracic radiographic findings (normality, cardiomegaly, alveolar pattern, bronchial pattern, interstitial pattern, mass, pleural effusion pneumothorax, and megaoesophagus): ResNet-50 (a specific type of convolutional neural network) was capable of identifying all the findings (except for bronchial and interstitial patterns) with an AUC above 0.8.

The application of deep-learning algorithms to computed tomography has yet not been explored in veterinary medicine.

Main complications reported for hepatic and splenic biopsy

Histological examination is, nowadays, the gold standard technique to obtain the final diagnosis of focal hepatic and splenic lesions [54,55]. Haemorrhage and thrombosis after cytological or tru-cut sampling procedures are reported to be common, especially when a neoplastic lesion is sampled [54,56].

The most common approach for focal splenic lesions, especially when large masses are present, is surgical splenectomy [27,57]. Nevertheless, complications following splenectomy (mainly haemorrhagic and thrombotic events), are reported in 7.6% of the patients [58]. On the other hand, cytology is reported to have a low sensitivity, but a high specificity, for both hepatic and splenic lesions [57,59]. As a consequence, when a cytological diagnosis of malignancy is obtained it is recomended to proceed in the a oncological and/or surgical treatment of the patient. Vice versa, when a cytological diagnosis of benign lesion is obtained, other diagnostic approaches might be suggested to confirm such a diagnosis.

AIM OF THE PhD PROJECT

Based on what discussed in the above paragraphs, the possibility to distinguish non-invasively between benign and malignant focal lesions might be a game changing approach in the diagnostic workflow of splenic and hepatic lesions.

The aim of this PhD project was to develop and train a deep learning-based algorithm capable of distinguishing between benign and malignant canine focal hepatic and splenic lesions based on their CT features. The entire study was composed by two parts: in the first part, we have developed a machine learning based algorithm based on the qualitative and quantitative CT features described by the radiologists. In the second part a CNN to predict whether a lesion was benign or malignant directly from the CT images was developed. Finally, a metanalysis study on focal liver lesions, with the aim to identify which of the qualitative and quantitative CT features reported in the literature are useful to distinguish malignant and benign hepatic lesions, was conducted.

BIBLIOGRAFIA

1. Withrow SJ, Page R, Vail D. *Withrow and MacEwen's Small Animal Clinical Oncology*. (Vail D, ed.). Elsevier; 2013. doi:10.1016/C2009-0-53135-2
2. Penninck, D.; d'Anjou M-A. *Atlas of Small Animal Ultrasonography*.; 2015.
3. Meuten DJ. *Tumors in Domestic Animals*.; 2017. doi:10.1002/9780470376928
4. Biswas M, Kuppili V, Saba L, et al. State-of-the-art review on deep learning in medical imaging. *Front Biosci - Landmark*. 2019;24(3):392-426. doi:10.2741/4725
5. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. *J Med Imaging Radiat Sci*. 2019;50(4):477-487. doi:10.1016/j.jmir.2019.09.005
6. Yang JD, Hainaut P, Gores GJ, et al. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*. 2019;16(10):589-604. doi:10.1038/s41575-019-0186-y.
7. Murakami T, Feeney DA, Bahr KL. Analysis of clinical and ultrasonographic data by use of logistic regression models for prediction of malignant versus benign causes of ultrasonographically detected focal liver lesions in dogs. *Am J Vet Res*. 2012;73(6):821-829. doi:10.2460/ajvr.73.6.821
8. O'Brien RT, Iani M, Matheson J, Delaney F, Young K. Contrast harmonic ultrasound of spontaneous liver nodules in 32 dogs. *Vet Radiol Ultrasound*. 2004;45(6):547-553. doi:10.1111/j.1740-8261.2004.04094.x
9. D'Onofrio M, Crosara S, De Robertis R, Canestrini S, Mucelli RP. Contrast-Enhanced Ultrasound of Focal Liver Lesions. *Am J Roentgenol*. 2015;205(1):W56-W66. doi:10.2214/AJR.14.14203
10. Sidhu PS, Cantisani V, Dietrich CF, et al. The EFSUMB guidelines and recommendations for the clinical practice of contrast-enhanced ultrasound (CEUS) in Non-Hepatic Applications: Update 2017 (Long Version). *Ultraschall der Medizin*. 2018;39(2):e2-e44. doi:10.1055/a-0586-1107
11. Chou R, Cuevas C, Fu R, et al. Imaging techniques for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis. *Ann Intern Med*. 2015;162(10):697-711. doi:10.7326/M14-2509
12. Nyman HT, Kristensen AT, Kjelgaard-Hansen M, McEvoy FJ. Contrast-enhanced ultrasonography in normal canine liver. Evaluation of imaging and safety parameters. *Vet Radiol Ultrasound*. 2005;46(3):243-250. doi:10.1111/j.1740-8261.2005.00034.x
13. Nakamura K, Takagi S, Sasaki N, et al. Contrast-enhanced ultrasonography for characterization

of canine focal liver lesions. *Vet Radiol Ultrasound*. 2010;51(1):79-85. doi:10.1111/j.1740-8261.2009.01627.x

14. Burti S, Zotti A, Rubini G, et al. Contrast-enhanced ultrasound features of malignant focal liver masses in dogs. *Sci Rep*. 2020;10(1):1-12. doi:10.1038/s41598-020-63220-3
15. Banzato T, Rubini G, Orlandi R, Bargellini P, Bonsembiante F, Zotti A. Contrast-enhanced ultrasound features of hepatocellular carcinoma in dogs. *Vet Rec*. 2019:1-9. doi:10.1136/vr.105282
16. Wisner ER, Zwingenberger AL. *Atlas of Small Animal CT and MRI*. first. Wiley-Blackwell; 2015.
17. Ariff B, Lloyd CR, Khan S, et al. Imaging of liver cancer. *World J Gastroenterol*. 2009;15(11):1289-1300. doi:10.3748/wjg.15.1289
18. Shah S, Shukla A, Paunipagar B. Radiological Features of Hepatocellular Carcinoma. *J Clin Exp Hepatol*. 2014;4(August):S63-S66. doi:10.1016/j.jceh.2014.06.009
19. Taniura T, Marukawa K, Yamada K, Hikasa Y, Ito K. Differential diagnosis of hepatic tumor-like lesions in dog by using dynamic CT scanning. *Hiroshima J Med Sci*. 2009;58(1):17-24.
20. Fukushima K, Kanemoto H, Ohno K, et al. Ct characteristics of primary hepatic mass lesions in dogs. *Vet Radiol Ultrasound*. 2012;53(3):252-257. doi:10.1111/j.1740-8261.2011.01917.x
21. Kutara K, Seki M, Ishikawa C, et al. Triple-phase helical computed tomography in dogs with hepatic masses. *Vet Radiol Ultrasound*. 2014;55(1):7-15. doi:10.1111/vru.12099
22. Griebie ER, David FH, Ober CP, et al. Evaluation of canine hepatic masses by use of triphasic computed tomography and B-mode, color flow, power, and pulsed-wave Doppler ultrasonography and correlation with histopathologic classification. *Am J Vet Res*. 2017;78(11):1273-1283.
23. Leela Arporn R, Ohta H, Shimbo G, et al. Computed tomographic features for differentiating benign from malignant liver lesions in dogs. *J Vet Med Sci*. 2019;81(12):1697-1704. doi:10.1292/jvms.19-0278
24. Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. *Front Vet Sci*. 2021;8(March):1-10. doi:10.3389/fvets.2021.611556
25. Stehlík L, Di Tommaso M, Signore F Del, et al. Triple-phase multidetector computed tomography in distinguishing canine hepatic lesions. *Animals*. 2021;11(1):1-12. doi:10.3390/ani11010011
26. Spangler WL, Culbertson MR. Prevalence, type, and importance of splenic diseases in dogs:

1,480 cases (1985-1989). *J Am Vet Med Assoc*. 1992;15(200(6)):829-834.

27. Cleveland MJ, Casale S. Incidence of malignancy and outcomes for dogs undergoing splenectomy for incidentally detected nonruptured splenic nodules or masses: 105 cases (2009-2013). *J Am Vet Med Assoc*. 2016;248(11):1267-1273. doi:10.2460/javma.248.11.1267
28. Boston SE, Higginson G, Monteith G. Concurrent splenic and right atrial mass at presentation in dogs with HSA: A retrospective study. *J Am Anim Hosp Assoc*. 2011;47(5):336-341. doi:10.5326/JAAHA-MS-5603
29. Schwarz LA, Penninck DG, Gliatto J. Ultrasound corner canine splenic myelolipomas. *Vet Radiol Ultrasound*. 2001;42(4):347-348. doi:10.1111/j.1740-8261.2001.tb00951.x
30. Ohlerth S, Dennler M, Ruefli E, et al. Contrast Harmonic Imaging Characterization of Canine Splenic Lesions. *J Vet Intern Med*. 2008;22:1095-1102. doi:10.1111/j.1939-1676.2008.0154
31. Nakamura K, Sasaki N, Murakami M, et al. Contrast-Enhanced Ultrasonography for Characterization of Focal Splenic Lesions in Dogs. *J Vet Intern Med*. 2010;24(6):1290-1297. doi:10.1111/j.1939-1676.2010.0609.x
32. Kutara K, Seki M, Ishigaki K, et al. Triple-phase helical computed tomography in dogs with solid splenic masses. *J Vet Med Sci*. 2017;79(11):1870-1877. doi:10.1292/jvms.17-0253
33. Jones ID, Lamb CR, Drees R, Priestnall SL, Mantis P. Associations between dual-phase computed tomography features and histopathologic diagnoses in 52 dogs with hepatic or splenic masses. *Vet Radiol Ultrasound*. 2016;57(2):144-153. doi:10.1111/vru.12336
34. Fife WD, Samii VF, Drost T, Mattoon JS, Hoshaw-Woodard S. Comparison between malignant and nonmalignant splenic masses in dogs using contrast-enhanced computed tomography. *Vet Radiol Ultrasound*. 2004;45(4):289-297. doi:10.1111/j.1740-8261.2004.04054.x
35. Cordella A, Caldin M, Bertolini G. Splenic extramedullary hematopoiesis in dogs is frequently detected on multiphase multidetector-row CT as hypervascular nodules. *Vet Radiol Ultrasound*. 2020;61(5):512-518. doi:10.1111/vru.12872
36. Rosenblatt F. *The Perceptron - A Perceiving and Recognizing Automaton*.; 1957.
37. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Peter Campbell J. Introduction to machine learning, neural networks, and deep learning. *Transl Vis Sci Technol*. 2020;9(2):1-12. doi:10.1167/tvst.9.2.14
38. Moawad AW, Fuentes DT, Elbanan MG, et al. Artificial Intelligence in Diagnostic Radiology: Where Do We Stand, Challenges, and Opportunities. *J Comput Assist Tomogr*. 2022;46(1):78-90. doi:10.1097/RCT.0000000000001247

39. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. doi:10.1038/nature14539

40. Yun SJ, Kim HC, Yang DM, Kim SW, Rhee SJ, Ahn SE. Diagnostic errors when interpreting abdominopelvic computed tomography: A pictorial review. *Br J Radiol*. 2017;90(1073):2-7. doi:10.1259/bjr.20160928

41. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction. *Radiographics*. 2015;35(6):1668-1676. doi:10.1148/rg.2015150023

42. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive error in radiology. *Am J Roentgenol*. 2017;208(4):739-749. doi:10.2214/AJR.16.16963

43. Siewert B, Sosna J, McNamara A, Raptopoulos V, Kruskal JB. Quality initiatives: Missed lesions at abdominal oncologic CT: Lessons learned from quality assurance. *Radiographics*. 2008;28(3):623-638. doi:10.1148/rg.283075188

44. Pinto A. Spectrum of diagnostic errors in radiology. *World J Radiol*. 2010;2(10):377. doi:10.4329/wjr.v2.i10.377

45. Kliewer MA, Brinkman MR, Hinshaw JL. The Back Alleys and Dark Corners of Abdomen and Pelvis Computed Tomography: The Most Frequent Sites of Missed Findings in the Multiplanar Era. *J Clin Imaging Sci*. 2020;10(70):70. doi:10.25259/jcis_184_2020

46. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: Current applications and future directions. *PLoS Med*. 2018;15(11):2-5. doi:10.1371/journal.pmed.1002707

47. Kakeda S, Moriya J, Sato H, et al. Improved Detection of Lung Nodules on Chest Radiographs Using a Commercial Computer-Aided Diagnosis System. *Am J Roentgenol*. 2004;182(2):505-510. doi:10.2214/ajr.182.2.1820505

48. Drost T, Henry G, Meinkoth JH, Woods P, Lehenbauer T. Quantification of hepatic and renal cortical echogenicity in clinically normal cats. *Am J Vet Res*. 2020;61:1016-1020.

49. Ivančić M, Mai W. Qualitative and quantitative comparison of renal vs. hepatic ultrasonographic intensity in healthy dogs. *Vet Radiol Ultrasound*. 2008;49(4):368-373. doi:10.1111/j.1740-8261.2008.00383.x

50. Banzato T, Bernardini M, Cherubini GB, Zotti A. Texture analysis of magnetic resonance images to predict histologic grade of meningiomas in dogs. *Am J Vet Res*. 2017;78(10):1156-1162. doi:10.2460/ajvr.78.10.1156

51. Marschner CB, Kokla M, Amigo JM, Rozanski EA, Wiinberg B, McEvoy FJ. Texture analysis of

pulmonary parenchymateous changes related to pulmonary thromboembolism in dogs - a novel approach using quantitative methods. *BMC Vet Res*. 2017;13(1):1-10. doi:10.1186/s12917-017-1117-1

52. Burti S, Longhin Osti V, Zotti A, Banzato T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *Vet J*. 2020;262:105505. doi:10.1016/j.tvjl.2020.105505
53. Banzato T, Wodzinski M, Burti S, et al. Automatic classification of canine thoracic radiographs using deep learning. *Sci Rep*. 2021;11(1):1-8. doi:10.1038/s41598-021-83515-3
54. Reece J, Pavlick M, Penninck DG, Webster CRL. Hemorrhage and complications associated with percutaneous ultrasound guided liver biopsy in dogs. *J Vet Intern Med*. 2020;34(6):2398-2404. doi:10.1111/jvim.15942
55. Lidbury JA. Getting the Most Out of Liver Biopsy. *Vet Clin North Am - Small Anim Pract*. 2017;47(3):569-583. doi:10.1016/j.cvsm.2016.11.007
56. McDevitt HL, Mayhew PD, Giuffrida MA, Brown DC, Culp WTN, Runge JJ. Short-term clinical outcome of laparoscopic liver biopsy in dogs: 106 cases (2003-2013). *J Am Vet Med Assoc*. 2016;248(1):83-90. doi:10.2460/javma.248.1.83
57. Tecilla M, Gambini M, Forlani A, Caniatti M, Ghisleni G, Roccabianca P. Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. *PLoS One*. 2019;14(11):1-15. doi:10.1371/journal.pone.0224945
58. Wendelburg KM, Toole TEO, McCobb E, et al. in Dogs Undergoing Splenectomy. *Javma*. 2014;245(Dec 15):1382-1390. doi:10.2460/javma.245.12.1382
59. Bahr KL, Sharkey LC, Murakami T, Feeney DA. Accuracy of US-guided FNA of focal liver lesions in dogs: 140 cases (2005-2008). *J Am Anim Hosp Assoc*. 2013;49(3):190-196. doi:10.5326/JAAHA-MS-5851

CHAPTER TWO

Diagnostic accuracy of delayed phase post contrast computed tomographic images in the diagnosis of focal liver lesions in dogs: 69 cases.

The chapter was adapted from:

Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. *Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases*. Front Vet Sci. 2021 Mar 4;8:611556. doi: 10.3389/fvets.2021.611556. Erratum in: Front Vet Sci. 2021 Nov 04;8:782672. PMID: 33748206; PMCID: PMC7969650.

ABSTRACT

To describe the computed tomographic (CT) features of focal liver lesions (FLLs) in dogs, that could enable predicting lesion histotype. Dogs diagnosed with FLLs through both CT and cytopathology and/or histopathology were retrospectively collected. Ten qualitative and 6 quantitative CT features have been described for each case. Lastly, a machine learning-based decision tree was developed to predict the lesion histotype. Four categories of FLLs - hepatocellular carcinoma (HCC, n=13), nodular hyperplasia (NH, n=19), other benign lesions (OBL, n=18), and other malignant lesions (OML, n=19) - were evaluated in 69 dogs. Five of the observed qualitative CT features resulted to be statistically significant in the distinction between the 4 categories: surface, appearance, lymph-node appearance, capsule formation, and homogeneity of contrast medium distribution. Three of the observed quantitative CT features were significantly different between the 4 categories: the Hounsfield Units (HU) of the radiologically normal liver parenchyma during the pre-contrast scan, the maximum dimension, and the ellipsoid volume of the lesion. Using the machine learning-based decision tree, it was possible to correctly classify NHs, OBLs, HCCs, and OMLs with an accuracy of 0.74, 0.88, 0.87, and 0.75 respectively. The developed decision tree could be an easy-to-use tool to predict the histotype of different FLLs in dogs. Cytology and histology are necessary to obtain the final diagnosis of the lesions.

INTRODUCTION

Focal liver lesions (FLLs) are common in dogs, especially in older patients. Malignant primary hepatic tumours represent less than 1.5% of all the malignant tumours in dog. More than 50% of malignant FLLs in dogs are hepatocellular carcinomas (HCC), whereas bile duct carcinomas (BDC) account for 22-40% of cases. Metastatic tumours (from pancreatic, splenic and gastroenteric tract in most cases) involving the liver are 2.5 times more frequent than primary tumours (2–4). Benign FLLs, especially in older dogs, are mostly nodular hyperplasia (NH), hepatocellular adenoma, and bile duct adenoma (2,5).

In human medicine the computed tomographic (CT) features of different FLLs in the arterial, portal and delayed phase are well described, and, therefore, it is possible to infer the histopathologic subtype of a FLL from its CT features (6–8). For example, the presence of a hypervascular pattern in a heterogeneous enhancing hepatic lesion during the arterial phase, is a feature often associated with HCCs (9).

In human medicine, the increasing availability of triple-phase CT, magnetic resonance imaging (MRI) and positron emission tomography (PET) have improved the scope to detect and diagnose FLLs (10). Moreover, the etiopathology of HCCs is well known (10), and the diagnostic and therapeutical approaches to this tumour are well detailed and based on the appearance of the lesions during the imaging examinations and on the histotype of the lesions (10,11). In animals, the etiopathology of FLLs is less detailed compared to human medicine, studies have involved a lower number of cases, and the inclusion and exclusion criteria related to the features of the lesions (e.g. margins, number, dimension) are less restrictive.

The CT features of FLLs in dogs have seldom been described (5,12–15). The studies on this topic are very heterogenous, indeed, a variety of features, scanning protocols, lesions, were evaluated in the different studies. A moderate to high accuracy of some CT features (e.g.: delayed phase enhancement (15,16)), in the distinction between benign and malignant masses is reported. Despite such encouraging results previous studies have only considered either only benign or malignant lesions (15,16) , or only certain lesion histotypes (5,12,13). Another important limitation is that, to date, most of the studies have only evaluated the accuracy of individual features and an algorithm to classify hepatic lesions based on their CT features has not been developed yet. In such a scenario, histopathology is still the gold-standard method for characterisation of FLLs in dogs.

The possible applications of machine learning algorithms have been widely explored in human diagnostic imaging in the last decades (17). On the other hand, in the last decade, the possibilities offered by this technology has raised an increasing interest also in veterinary medicine (18–22). Machine learning comprise a wide range of algorithms that can be broadly divided into machine learning and deep learning (23). Generally speaking, the main difference between the two systems is that machine learning algorithms are statistical methods that are applied on human crafted features extracted from the images whereas deep learning algorithms, are end-to-end algorithms capable of automatically extracting the features from the images and analysing them accordingly.

The aims of the present study are: 1) to describe the quantitative and qualitative computed tomographic features in pre-contrast and in the delayed phase of different histopathological subtypes of FLLs in dogs; 2) to develop a machine learning-based decision tree to assist the radiologist and the clinician in predicting different histopathological subtypes of FLLs based on their CT features. We hypothesized that a diagnostic algorithm, based on the CT features as described by the radiologist, could help in the prediction of the histopathological type of FLL.

MATERIAL AND METHODS

Study Population

Dogs referred to the Pedrani Veterinary Clinic (Via Caldierino 14, Zugliano, Vicenza, Italy) and to the Veterinary Teaching Hospital of the University of Padua (Viale Dell’Università 16, Legnaro, Padua, Italy) between June 2015 and January 2020 and which underwent computed tomographic examination and had FLL diagnosed with cytopathology and/or histopathology were retrospectively selected. Complete signalment was recorded for each patient. Dogs with hepatic masses/nodules not diagnosed on pathology, or which underwent chemotherapy before the tomographic examination, were excluded.

All the methods were carried out in accordance with the relevant guidelines and regulations. This study was conducted respecting the Italian law D. Leg.vo 26/2014 (that transposes EU Directive 2010/63/EU). Nevertheless, since the data used in this study were part of routine clinical activity, no ethical committee approval was required. Informed consent regarding the treatment of personal data was obtained from the owners.

Cytopathological and Histopathological Examination

Cytological samples were obtained through ultrasound-guided fine-needle aspiration of the hepatic mass. 21 Gauge needles were always used. Aspirates were spread on glass slides that were air-dried, stained with May-Grünwald-Giemsa stain and cover-slipped. The evaluation of the cytological slides was performed by one cytologist (FB).

Histological samples were obtained through ultrasound-guided Tru-cut biopsy of the hepatic mass. Tissue samples were fixed in 10% neutral formalin, processed by dehydration in a graded ethanol series and embedded in paraffin. Histological examination was carried out on 4-µm-thick sections stained with haematoxylin and eosin by one pathologist.

Computed Tomography Examination

All the animals were fasted for a 12-hour-period prior to examination. General anaesthesia was always administered. The CT examinations were performed using three different scanners (Asteion super4, Toshiba Medical System Corporation; Revolution ACT, General Electric Medical System; Optima CT 520 Series, General Electric Medical System). Due to the different technology of the scanners, slightly different scanning protocols were used. The scanning protocols were as follows: In both facilities contrast medium (Ioversol 350mg/ml, Optiray 350, LIEBEL-FLARSHEIM COMPANY LLC, USA) at the dosage of 660mg I/kg of body weight was injected through an IV catheter placed in the cephalic vein. At the Pedrani Veterinary Clinic the contrast medium was administered by means of an injector. At the Veterinary Teaching Hospital, the contrast medium was manually administered through an intravenous bolus at the fastest possible rate. In both institutions a standard total-body scan with a pre-contrast and a delayed phase, starting from the nose tip at 60-70 seconds after the end of contrast medium injection, was used. This means that, considering the post-start injection-scanning at the liver site a delay ranging from 69 to 105 seconds at the Pedrani Veterinary Clinic and a 74 to 120 seconds at the Veterinary teaching Hospital should be considered. All the patients were placed on ventral recumbency during the scan.

All the images were stored as digital imaging and communication in medicine (DICOM) files.

Image Analysis

All the scans were reviewed by two experienced radiologists (AZ and SB) using a picture archiving and communication system (PACS) workstation (RadiAnt DICOM Viewer 5.5.0). The qualitative and quantitative CT features were evaluated during both the pre-contrast phase and the delayed phase. In the case of multiple lesions, only the CT features of the lesions that were sampled have been described. All the studies were displayed in a soft tissue window (WW: 400 HU– WL: 40HU)

The following qualitative features were evaluated: 1) margins (well- or ill-defined); 2) surface (regular or irregular); 3) appearance (solid or cyst-like) - the lesion was classified as “cyst-like” if at least one area having a measured Hounsfield Unit (HU) value similar to that of the gallbladder of the same animal (representing possible necrosis or haemorrhage), was present (12); 4) portal lymph-nodes appearance (normal or abnormal) – portal lymph nodes were graded as abnormal if any of the following changes were evident: a) lymphoadenomagaly b) heterogeneous c) irregular shape; 5) capsule formation (present or absent) - the presence of a capsule was reported if a thin and hyperenhancing fibrous peripheral border, encompassing most of the lesion, was present and care was placed to differentiate between real capsule formation and the presence of enlarged vessels or sinusoids mimicking the presence of a real capsule (24); 6) portal invasion, meaning the invasion of the portal vein and its branches (present or absent); 7) homogeneity in the distribution of the contrast medium inside the lesion (homogeneous or heterogeneous); 8) enhancement pattern (prevalently central, marginal, or diffuse distribution).

The following quantitative characteristics were evaluated: 1) attenuation (measured as an HU value) of the radiographically normal liver parenchyma surrounding the lesion, both in pre- and post-contrast scans; 2) attenuation (HU value) of the lesion both in pre- and post-contrast scans; 3) maximum transverse diameter; 4) volume - the shape of the lesion was considered to be an ellipsoid and the formula $V = \frac{4}{3}\pi\ (\mathbf{height} * \mathbf{width} * \mathbf{length})$ was applied (25); 5) attenuation of the lesion compared to that of the radiologically normal liver parenchyma in the pre-contrast images (hypoattenuating, isoattenuating or hyperattenuating); 6) enhancement degree of the lesion compared to that of the radiologically normal liver parenchyma in post-contrast images (hypoenhancing, isoenhancing or hyperenhancing). The attenuation and the enhancement degree of the lesion were determined based on the difference between the HU value measured in the lesion and the HU value measured on the radiologically normal liver parenchyma. If the difference between the lesion and the parenchyma felt in the ±10 HU range, the lesion was classified as iso-attenuating/enhancing; with a difference greater than +10 HU the lesion was classified as hyper-attenuating/enhancing; if the difference was lower than -10 HU, the lesion was classified as hypo-attenuating/enhancing (13,14). The CT features were evaluated by two of the authors of this study (SB: with 4 years of experience in diagnostic imaging and AZ: with 20 years of experience in diagnostic imaging) that were blinded to the results of the histopathological examination.

HU values were measured on the three circular regions of interest (ROIs) which could be placed in different regions of both normal and pathological parenchyma. The same ROIs were placed in pre- and post-contrast images.

Statistical Analysis

All the statistical analysis was performed using R-software version 3.6.1 (R Core Team (2019); R: a language and environment for statistical computing. The R Foundation for Statistical Computing, Vienna, Austria). The difference in the distribution of the qualitative variables was analysed with the chi-Square test (χ^2) or with Fisher's exact method. Post-hoc multiple comparisons among levels were performed using Marascuilo approach. Differences in the distribution of the quantitative variables were analysed with a one-way analysis of variance (ANOVA) for normally distributed data, or with the Kruskal-Wallis test for non-normally distributed data. The Tukey-Kramer method was used for multiple comparison tests after ANOVA analysis. A Steel-Dwass-Critchlow-Fligner procedure was used for pairwise comparison testing after Kruskal-Wallis analysis. A p-value of < 0.05 was considered as statistically significant.

A machine learning-classification tree analysis was performed to detect the best discriminating CT features. A recursive partitioning method was adopted using the rpart package of R(26). This package builds a decision tree based on a three-step procedure. In the first step, the feature that provides the best splitting of the data into two groups is selected. The second step of the procedure uses a 10-fold cross-validation to select the tree having both the lowest number of branches and the lowest misclassification rate. Thereafter, the developed tree is applied to the original dataset and sensibility, specificity, accuracy and misclassification rate are calculated.

RESULTS

Patients

Based on the inclusion criteria, 69 dogs of different breeds (37 females and 32 males, with mean age of 11 years ranging from 4 to 16.5 years), with pathologically diagnosed FLLs, and which underwent a CT examination, were included. Cytopathology was performed in 54 dogs. Tru-cut biopsy was performed in 13 dogs, and both cytology and histology were performed in 3 dogs. In two of these latter cases there was an agreement between cytology and histology (both were suggestive of HCC); in one case the cytological diagnosis was blood collection and vacuolar degeneration, while NH was diagnosed by means of histology. Benign lesions were diagnosed in 37 cases (1 biliary duct adenoma, 1 haematoma, 1 inflammation, 2 hepatocellular adenomas, 2 normal liver parenchyma, 11 degenerations, and 19 nodular hyperplasia), and malignant lesions were diagnosed in 32 cases (1 mast cell tumour, 1 plasmocytoma, 1 biliary duct carcinoma, 1 undifferentiated carcinoma, 1 melanoma, 1 metastasis of mammary neoplasia, 2 lymphomas, 4 endocrine neoplasia, 7 sarcomas and 13 HCCs).

Due to the large variability in the histological subtypes of the lesions included in the study, the patients were grouped into the following four categories for the statistical analysis: NHs (19 cases); OBLs (18 cases), HCCs (13 cases), and OMLs (19 cases).

Image Analysis

A summary of all the CT parameters evaluated, along with the p-values of the statistical tests, is reported in Table 1 (qualitative features) and Table 2 (quantitative features).

Table 1 Number of cases, classified based on cytological or histological examination, showing the qualitative features, along with the p-value.

	Category					
	Nodular hyperplasia	**Other benign lesions**	**Hepatocarcinoma**	**Other malignant lesions**	**Total**	**p-value**
	(n = 19)	*(n = 18)*	*(n = 13)*	*(n = 19)*	*(n = 69)*	
Margins						0.07
well defined	14 (73.7%)	13 (72.2%)	13 (100%)	11 (57.9%)	51	
ill defined	5 (26.3%)	5 (27.8%)	0	8 (47.1%)	18	
Surface						< 0.01
regular	5 (26.3%)	11 (61.1%)	0	2 (10.5%)	18	
irregular	14 (73.7%)	7 (38.9%)	13 (100%)	17 (89.5%)	51	
Aspect						0.03
solid	8 (42.1%)	9 (50.0%)	1 (7.7%)	9 (47.4%)	27	
cyst-like	11 (57.9%)	9 (50.0%)	12 (92.3%)	10 (52.6%)	42	
Lymph-nodes						< 0.01
normal	14 (73.7%)	17 (94.4%)	6 (46.2%)	8 (42.1%)	45	
abnormal	5 (26.3%)	1 (5.6%)	7 (53.8%)	11 (57.9%)	24	
Portal invasion						0.39
present	0	0	0	2 (10.5%)	2	
absent	19 (100%)	18 (100%)	13 (100%)	17 (89.5%)	67	
Capsule formation						0.03
present	0	7 (38.9%)	7 (53.8%)	9 (47.4%)	23	
absent	19 (100%)	11 (61.1%)	6 (46.2%)	10 (52.6%)	46	
Homogeneity post-contrast medium						< 0.01
homogeneous	7 (36.8%)	12 (66.7%)	0	6 (31.6%)	25	
heterogeneous	12 (63.2%)	6 (33.3%)	13 (100%)	13 (68.4%)	54	
Enhacement pattern						0.70
diffuse	16 (84.2%)	14 (77.8%)	11 (84.6%)	15 (78.9%)	56	
marginal	3 (15.8%)	4 (22.2%)	2 (15.4%)	3 (15.8%)	12	
central	0	0	0	1 (5.3%)	1	
Attenuation pre-contrast medium						0.53
hypoattenuating	13 (68.4%)	15 (83.3%)	12 (92.3%)	16 (84.2%)	56	
isoattenuating	6 (31.6%)	3 (16.7%)	1 (7.7%)	3 (15.8%)	13	
hyperattenuating	0	0	0	0		
Enhancement post-contrast medium						0.65
hypoenhancing	15 (78.9%)	16 (88.9%)	13 (100%)	17 (89.5%)	61	
isoenhancing	1 (5.3%)	0	0	0	1	
hyperenhancing	3 (15.8%)	2 (11.1%)	0	2 (10.5%)	7	

Table 2 Quantitative features of the lesions, classified based on cytological or histological examination, are reported as medians along with the first and third quartile values and the p-value.

	Category				
	Nodular hyperplasia	**Other benign lesions**	**Hepatocarcinoma**	**Other malignant lesions**	**p-value**
	(n = 19)	*(n = 18)*	*(n = 13)*	*(n = 19)*	
HU normal liver pre-contrast medium	63.82 (53.79-69.79)^{ab}	66.84 (64.36-72.54)^{a}	58.63 (53.12-63.02)^{b}	60.03 (54.59-64.55)^{b}	< 0.01
HU normal liver post-contrast medium	144.54 (120.59-169.15)	137.6 (126.71-154.01)	127.72 (116.12-135.06)	142.87 (117.56-157.18)	0.29
HU lesion pre-contrast medium	45.68 (40.72-54.79)	39.5 (29.94-45.99)	41.48 (34.87-46.93)	39.93 (34.39-46.12)	0.8
HU lesion post-contrast medium	114.37 (50.96-144.87)	75.65 (61.37-121.17)	67.39 (56.03-83.93)	83.19 (66.32-121.40)	0.13
Max dimension	4.53 (2.45-6.75)^{ab}	2.15 (1.12-5.33)^{b}	11.11 (5.67-13.76)^{a}	3.59 (2.11-4.61)^{b}	< 0.01
Ellipsoid volume	40.78 (6.15-112.86)^{ab}	2.41 (0.39-26.78)^{c}	393.57 (54.80-727.31)^{a}	8.31 (3.67-23.60)^{bc}	< 0.01

Among the qualitative features the surface ($\chi^2 = 19.80$; p-value < 0.01), appearance ($\chi^2 = 8.75$; p-value = 0.03), lymph-node appearance ($\chi^2 = 13.19$; p-value < 0.01), capsule formation ($\chi^2 = 9.23$; p-value = 0.03), and homogeneity in the distribution of the contrast medium inside the lesion ($\chi^2 = 13.79$; p-value < 0.01) showed statistically significant differences. No significant differences were evident for the characteristics of the margins ($\chi^2 = 7.14$; p-value = 0.07), the presence of portal invasion ($\chi^2 = 3.02$; p-value = 0.39), the enhancement pattern ($\chi^2 = 3.83$; p-value = 0.70), attenuation during the pre-contrast phase ($\chi^2 = 2.21$; p-value = 0.53), and enhancement during the post-contrast phase ($\chi^2 = 4.19$; p-value = 0.65). Most of the HCCs included in this study (>90%), showed an irregular surface, a cyst-like appearance and heterogeneity post-contrast medium. The presence of abnormal lymph-nodes was equally recorded both in HCC and OML (more than 50% of the cases). The presence of a capsule was recorded in all lesions, except NH.

All the quantitative variables showed non-normal distribution. Therefore, values are reported as medians along with the first and third quartiles, and differences were calculated using the Kruskal-Wallis test. As a result of the Kruskal-Wallis test, the HU value of the radiologically normal liver parenchyma during the pre-contrast phase (K = 12.71; p-value < 0.01), maximum dimension (K = 14.60; p-value < 0.01) and ellipsoid volume (K = 18.21; p-value < 0.01) showed statistically significant differences. Post-hoc testing revealed the following significant differences: 1) between HCCs and OBLs regarding maximum dimension (p-value < 0.01), ellipsoid volume (p-value < 0.01) and the HU value of the radiologically normal liver parenchyma during the pre-contrast phase (p-value < 0.01); 2) between HCCs and OMLs regarding maximum dimension (p = 0.03), and ellipsoid volume (p-value < 0.01); 3) between nodular hyperplasia and OBLs (p-value = 0.04); 4) between other malignant and OBLs for the HU value of the radiologically normal liver parenchyma during the pre-contrast phase (p-value < 0.01).

No statistically significant differences were evident for the remaining quantitative features: the HU value of the radiologically normal liver parenchyma during the post-contrast phase (K = 3.80; p-value = 0.29), and the HU value of the lesions during both pre-contrast (K = 2.76; p-value = 0.43) and post-contrast (K = 4.93; p-value = 0.18) phases.

NH is a benign condition, which is often an occasional finding, but it could be a diagnostic challenge in a patient with a known history for another malignancy (24). In the present study, NHs were mainly hypoattenuating (68.4%) and hypoenhancing (78.9%) lesions showing a diffuse contrast enhancement pattern (84.2%), with homogeneous distribution (63.2%), well-defined margins

(73.7%), an irregular surface (73.7%) and a cyst-like appearance (57.9%). The hepatic lymph nodes were radiologically normal in 14 out of 19 cases (73.7%), whereas neither portal vein invasion nor a real capsule formation were evident in any case.

The OBLs were mainly hypoattenuating (83.3%) and hypoenhancing (88.9%) lesions characterised by a diffuse enhancement pattern (77.8%), homogeneous distribution (66.7%), well-defined margins (72.2%), a regular surface (61.1%), and with both a cyst-like (50%) and solid (50%) appearance. Hepatic lymph nodes were almost always radiologically normal (17/18), and portal vein invasion was never detected. A fibrous capsule was evident in 7 out of 18 patients.

HCCs were mostly cyst-like (12/13, 92.3%), hypoattenuating (12/13) and hypoenhancing lesions (100%), with a diffuse enhancement pattern (84.6%), a heterogeneous distribution (100%), well-defined margins (100%), and an irregular surface (100%). The hepatic lymph nodes were abnormal in 53.8% of the cases (7/13), and portal vein invasion was never evident. The lesions were surrounded by a fibrous capsule in 53.8% of cases. Post-hoc tests revealed no significant differences between HCCs and NHs for both qualitative variables and quantitative variables.

The OMLs were mainly hypoattenuating (84.2%) and hypoenhancing (89.5%), showing a diffuse enhancement pattern (78.9%), heterogeneous distribution (68.4%), well-defined margins (57.9%) and an irregular surface (89.5%). They had both a cyst-like (52.6%) and a solid (47.4%) appearance. The hepatic lymph nodes were abnormal in 57.9% of cases. Portal vein invasion was evident only in 2 patients (10.5%). Fibrous capsule formation was evident in 9 out of 19 (47.4%).

Representative cases of each FLL showing the most typical CT features are reported in Figures 1-4.

The final decision tree algorithm was built on five automatically selected CT features: 1 qualitative feature (lymph nodes), and 4 quantitative features (maximum dimension, HU normal liver pre-contrast, HU normal liver post-contrast, HU lesion pre-contrast). The confusion matrix is reported in Table 3. The sensitivity, specificity, accuracy, precision, Mathew's correlation coefficient, and the misclassification rate for each group, along with the global misclassification rate are reported in Table 4. The decision tree is set out in Figure 5.

The decision tree misclassified 7 cases as OBL and 5 cases as OML. Among the 7 incorrectly classified OBLs, 6 were in fact degenerations (misclassified as OML in 4 cases, and as NH in 2 cases). The remaining case classified as OBL was in fact a hepatocellular adenoma (misclassified as NH). Among the 5 incorrectly classified OMLs, 1 metastatic splenic sarcoma was misclassified as HCC. Lastly, 1 mast cell tumour, 1 endocrine neoplasia, 1 myeloma, and 1 metastasis of splenic sarcoma were classified as NH.

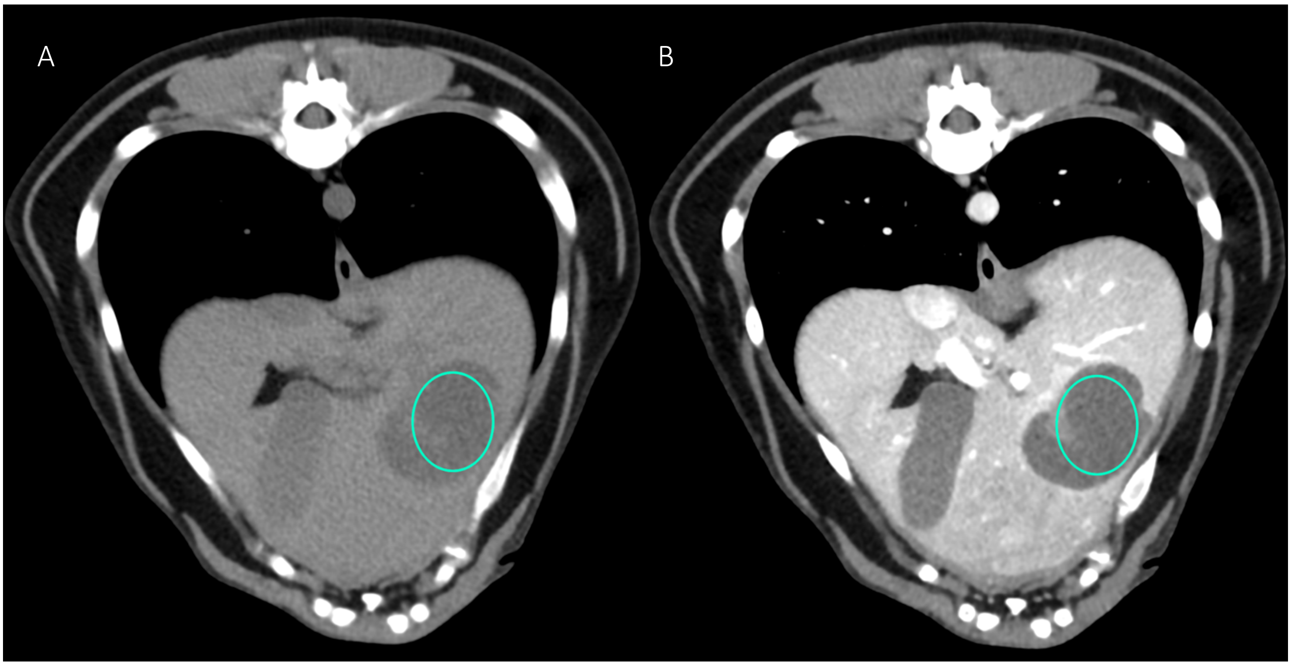


Figure 1 Example of a NH lesion that shows hypoattenuation and hypoenhancement, diffuse contrast enhancement pattern, with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. A. image obtained from the pre-contrast scan; B. image obtained from the delayed scan. A ROI is placed inside the lesion in both. Based on the developed decision tree this lesion was classified as OBL.

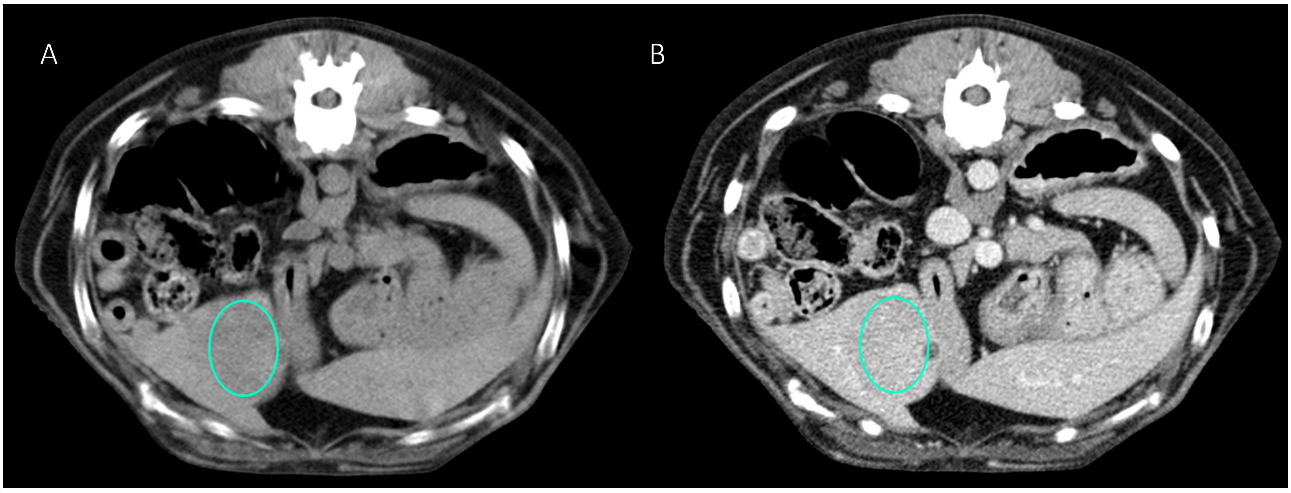


Figure 2 Example of an OBL lesion (diagnosed as adenoma) showing hyperattenuation and hypoenhancement, diffuse contrast enhancement pattern, with homogeneous distribution, well-defined margins, regular surface, and cyst-like appearance. A. image obtained from the pre-contrast scan; B. image obtained from the delayed scan. A ROI is placed inside the lesion in both. Based on the developed decision tree this lesion was classified as OBL.

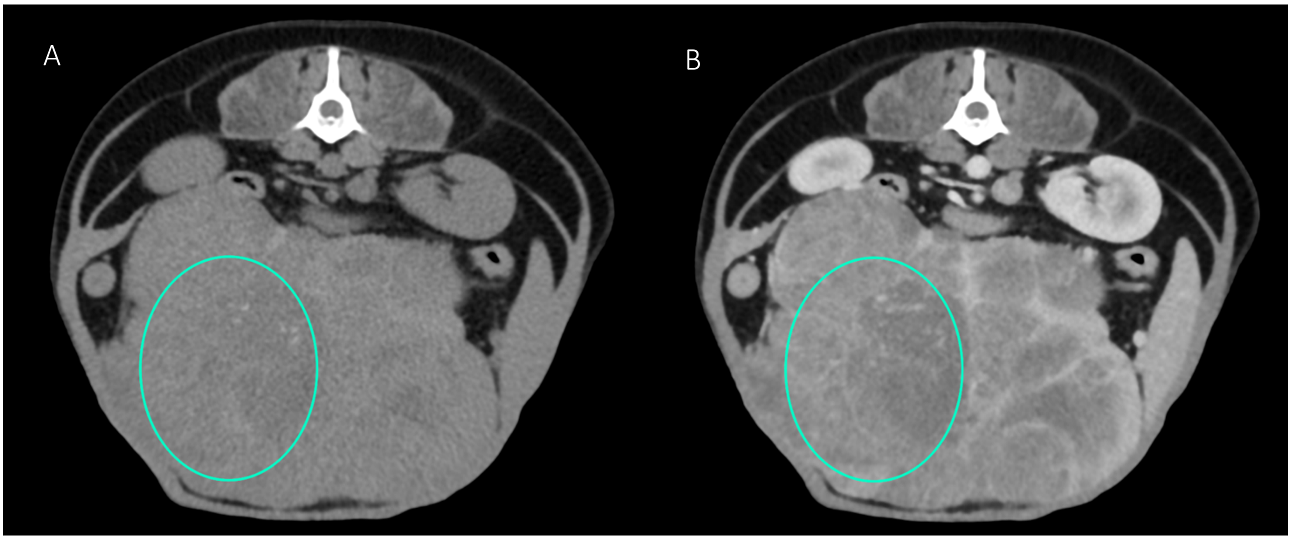


Figure 3 Example of an HCC showing hypoattenuation and hypoenhancement, diffuse contrast enhancement pattern, with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. A. image obtained from the pre-contrast scan; B. image obtained from the delayed scan. A ROI is placed inside the lesion in both. Based on the developed decision tree this lesion was classified as HCC.

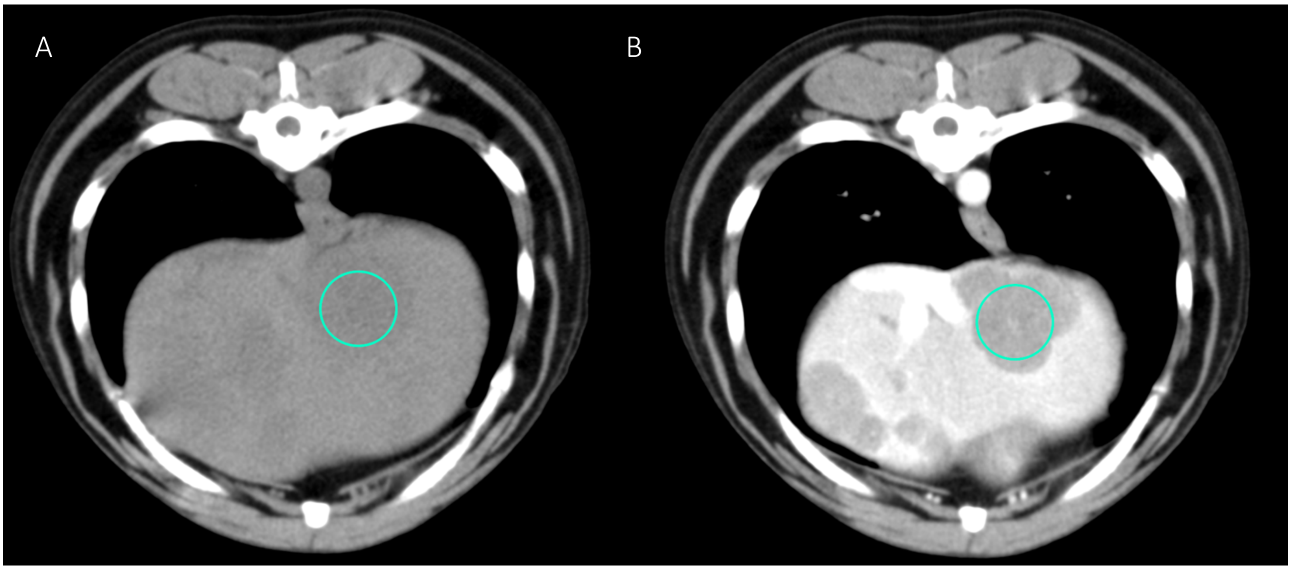


Figure 4 Example of an OML (diagnosed as lymphoma) showing hypoattenuation and hypoenhancement, diffuse contrast enhancement pattern, with heterogeneous distribution, well-defined margins, irregular surface, and solid appearance. A. image obtained from the pre-contrast scan; B. image obtained from the delayed scan. A ROI is placed inside the lesion in both. Based on the developed decision tree this lesion was classified as OML.

Table 3 Confusion matrix to summarize the performance of the machine learning algorithm giving the number of predicted cases among the four categories of FLLs.

		Actual				
		Nodular hyperplasia	**Other benign lesions**	**Hepatocarcinoma**	**Other malignant lesions**	**Total**
Predicted	**Nodular hyperplasia**	10	3	2	4	19
	Other benign lesions	1	11	0	0	12
	Hepatocarcinoma	3	0	8	1	12
	Other malignant lesions	5	4	3	14	26
	Total	19	18	13	19	69

Table 4 Complete results of the classification of the FLLs based on the machine learning-based decision tree.

	Nodular hyperplasia	**Other benign lesions**	**Hepatocarcinoma**	**Other malignant lesions**
Sensitivity (%)	53	61	62	74
Specificity (%)	82	98	93	76
Accuracy (%)	74	88	87	75
Precision (%)	53	92	67	54
MCC*	0.42	0.69	0.58	0.51
Miscassification rate (%)	47	39	38	26
Global misclassification rate (%)		38		

**MCC, Matthews Correlation Coefficient*

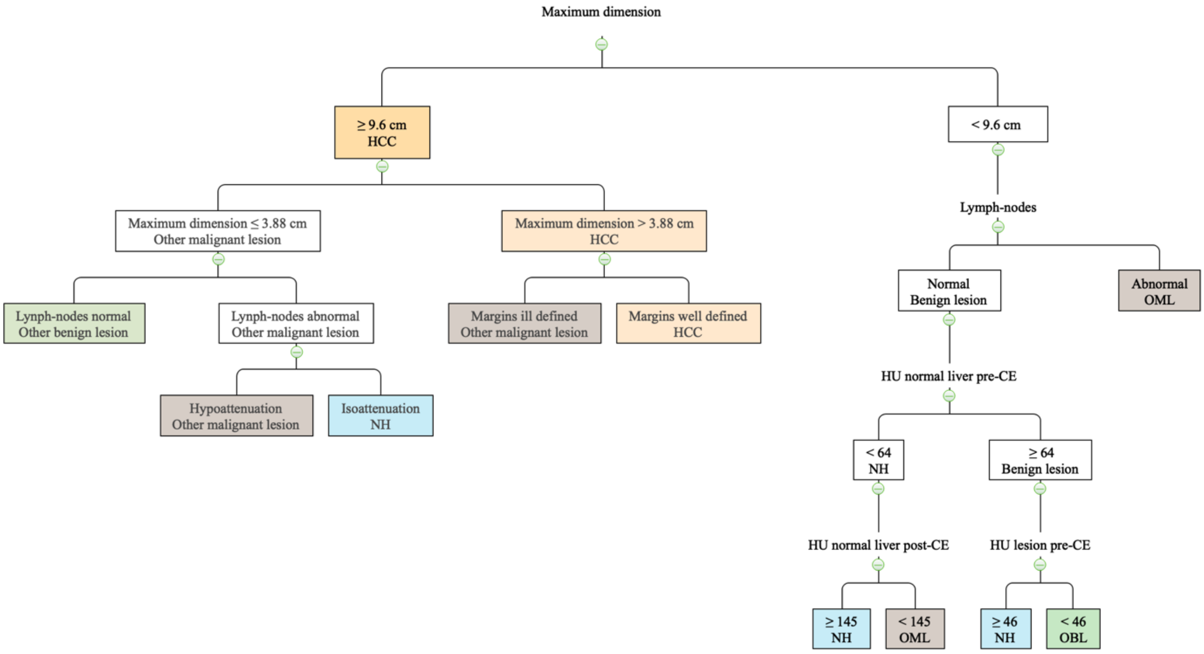


Figure 5 The machine learning-based decision tree developed based on qualitative and quantitative CT features of the FLLs.

DISCUSSION

A machine learning-based, easy-to-follow, algorithm to predict the histotype of canine liver lesions, based on their CT features, is proposed. The proposed algorithm had a variable accuracy in the classification of the different histotypes, ranging from 0.74 of NH to 0.88 of OBL. On the other hand, also the precision was extremely variable ranging from 0.53 of NH to 0.92 of OBL. The global misclassification rate was high with 38% of the lesion that were incorrectly classified. The same machine learning algorithm (decision tree) used in this paper was used by the authors in a previous report (3) on the CEUS features of canine malignant liver lesions. In that report, a higher number of cases (185 total cases) was available, thus offering the scope to split the dataset into a training and a test set. In the present study, the relatively low number of available cases did not allow such a division of the dataset, and, therefore, the accuracy of the decision tree was retested on the same dataset.

A machine learning-based approach was used by other authors (15). In their report, Griebie and others 2017 used both a stepwise discriminant analysis and a Fisher prediction equation to identify the CT or ultrasound features that might be helpful in distinguishing between focal liver lesions. Using such a classification method allowed them to accurately (sensitivity = 96.7; specificity = 87.5) distinguish only between benign and malignant lesions. On the other hand, none of the features described by Griebie and others 2017 resulted as significant when a specific diagnosis classification was used.

In the present study, the different histotypes of the lesions were grouped into four different categories, in order to have a sufficient number of cases in each category. Using such a classification scheme and applying the decision tree enabled us to progressively detect those features that might be helpful in distinguishing between specific categories of lesion. On the other contrary, using such an approach results in an overall lower accuracy of the model when compared to the binary classification proposed by Griebie and others 2017.

A limitation of the present study is that, due to the fact that the masses were mostly incidental findings, only the delayed phase was evaluated in this study. In other studies on the same topic(5,13,15,16,27) the enhancement patterns have been evaluated in triple phase (arterial, portal, delayed). On the other hand, most (not all) of the CT features that are reported to be helpful in the distinction among different FLL are seen in the delayed phase. There is no agreement among different authors regarding the CT features of canine HCCs in the delayed phase. Indeed, Taniura and others 2009, and Fukushima and others 2012 describe HCCs as hypoenhancing lesions, Jones and others 2016 as hyperenhancing, and Kutara and others 2014 report that HCCs might have all the possible enhancement patterns. The results of the present study were similar to the findings by Taniura and others 2009 and Fukushima and others 2012. Indeed, all the HCCs included in this study were hypoenhancing lesions. Furthermore, HCCs have been described as being cyst-like lesions (12) with a heterogeneous contrast enhancement (13). These finding were also similar to those reported the present study; indeed, 12/13 HCCs were cyst-like lesions and a heterogeneous distribution of the contrast medium was always evident. On the other hand, as a result of the decision tree, the most effective CT feature to classify HCCs was the maximum dimension of the lesion. Interestingly, the cut-off value identified by the decision tree (9.6 cm) was very similar to the cut-off value reported by Griebie and others 2017 (9.5 cm) to identify malignant lesions.

The presence of an hyperenhancing fibrous capsule surrounding the lesion is reported to be a distinctive CT feature of HCCs in humans, and, therefore, is used in the distinction between HCCs and NHs (24). While this feature is distinctive in people, the feature was only present in 7/13 dogs with HCC and in none of the NHs in the current study. Furthermore, also Taniura and others 2009 and Fukushima and others 2012 identified a hyperenhancing capsule in 25/36 and 13/14 HCC cases, respectively. Lastly, the presence of a hyperenhancing capsule in NHs in dogs is reported only in a single case (12). Nevertheless, the presence of such a capsule was also evident in 7/18 OBLs and in 9/19 OMLs.

There is also no agreement among different authors regarding the CT features of NH. Indeed, they are described as isoenhancing lesions in the venous phase by Taniura and others 2009, and Fukushima and others 2012. On the contrary, Kutara and others 2014 describe them as often being hyperenhancing or isoenhancing lesions showing a homogeneous distribution of the contrast medium. Interestingly, in the present study, most of the NH were hypoenhancing lesions showing a diffuse enhancement pattern. In human patients, NHs are reported to be mainly isoenhancing during the venous phase (24).

A limitation of the present study is that Tru-cut biopsies were obtained only in 13 patients and only 3 patients had both cytology and histology. Tru-cut and incisional biopsies are, nowadays, considered the gold-standard diagnostic techniques to determine the histopathological subtype of FLLs (28,29) and, indeed either Tru-cut (12,15) or surgical biopsies (5,13,16) have been used as reference standard in previous studies. However, because of the possible side effects, such as haemorrhage, hypercoagulable states, hypotension, peritonitis, hepatic emphysema (30–32), biopsies are not always performed in patients with FLL. In these cases, cytological examination represents a viable alternative to biopsies, even if the sensitivity for malignancy is lower (33,34). We are aware of the lack of standardization of the injection procedures and also those manual procedures are slightly operator-dependent; however, the key-point is that each subject of our study was scanned during an earlier or slightly more advanced stage of the delayed phase. We would like to state again that this is not a relevant hepatic phase study; in fact, a specific arterial or portal phase was not performed in any subject. This study reports an analysis of the hepatic pathological CT patterns within slightly variable stages of the delayed phase that could be found during non-focused CT whole body examinations.

Another possible limitation of the present study is that the CT features of the lesions were evaluated only during the delayed phase. This was carried out because of the different technology of the CT scanners used to acquire the images for this study. Indeed, both 4- and 16-row CT scanners were used. The former does not enable images to be acquired during the arterial phase and, therefore, in order to make a reliable comparison, only the delayed phase was analysed. At this point it is, however, important to stress that, in the study by Griebie and others 2017, using a multiphase CT scanner, only the CT features of the lesions in the venous and delayed phases were statistically significant for the development of their prediction model. Moreover, in a standard clinical protocol for CT total-body scan only the arterial and delayed phase are performed.

The decision tree, based on the qualitative and quantitative CT features of the lesions, reported in the present results could be an easy-to-use tool for the veterinary clinician in predicting the histotype of different canine FLLs. A larger number of cases, enabling application of stricter inclusion/exclusion criteria (for example using cut-off values for the dimensions of the lesions to be included) could, prospectively, enable creation of a more accurate decision tree. Nevertheless, as is often the case also in human medicine, the final histotype of a FLL should always be determined based on cytology or histology.

CONCLUSIONS

The CT features of 69 FLL, analyzed in the pre-contrast and in the delayed phase, are reported. The developed machine learning algorithm had a 62% overall accuracy in the classification of the FLL based on their CT features. The misclassification rate was highest (47%) for NH and lowest for OML (26%). The use of the proposed decision tree, could, prospectively help the clinician in the evaluation of FLL.

Funding The present paper is part of a project funded by a research grant from the Department of Animal Medicine, Production and Health – MAPS, University of Padua, Italy: SID- Zotti 2018 (€ 32,000; Application of deep-learning algorithms in pet animal diagnostic imaging)

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Patient consent for publication Not required.

Ethics approval This study was conducted according to the Italian law D. Leg.vo 26/2014 (that transposes the EU directive 2010/63/EU). As the data used in this study were part of the routine clinical activity no ethical committee approval was needed. Informed consent regarding the treatment of personal data was obtained from the owners.

BIBLIOGRAPHY

1. Itri JN, Tappouni RR, Mceachern RO, Pesch AJ, Patel SH. PATIENT-CENTERED CARE Fundamentals of Diagnostic Error in Imaging. RadioGraphics [Internet]. 2018;38(5):1845–65. Available from: https://doi.org/10.1148/rg.2018180021
2. Withrow SJ. Small Animal Clinical Oncology [Internet]. 2016. Available from: https://www.taylorfrancis.com/books/9781482225426
3. Banzato T, Rubini G, Orlandi R, Bargellini P, Bonsembiante F, Zotti A. Contrast-enhanced ultrasound features of hepatocellular carcinoma in dogs. Vet Rec. 2019;1–9.
4. Burti S, Zotti A, Rubini G, Orlandi R, Bargellini P, Bonsembiante F, et al. Contrast-enhanced ultrasound features of malignant focal liver masses in dogs. Sci Rep. 2020;10(1):1–12.
5. Taniura T, Marukawa K, Yamada K, Hikasa Y, Ito K. Differential diagnosis of hepatic tumor-like lesions in dog by using dynamic CT scanning. Hiroshima J Med Sci. 2009;58(1):17–24.
6. Nino-Murcia M, Olcott EW, Jeffrey RB, Lamm RL, Beaulieu CF, Jain KA. Focal liver lesions: Pattern-based classification scheme for enhancement at arterial CT. Radiology. 2000;215(3):746–51.
7. Lee KHY, O'Malley ME, Haider MA, Hanbidge A. Triple-Phase MDCT of Hepatocellular Carcinoma. Am J Roentgenol. 2004;182(3):643–9.
8. Song S, Li Z, Niu L, Zhou X, Wang G, Gao Y, et al. Hypervascular hepatic focal lesions on dynamic contrast-enhanced CT: preliminary data from arterial phase scans texture analysis for classification. Clin Radiol. 2019;74(8):653.e11-653.e18.
9. Lee KHY, O'Malley ME, Haider MA, Hanbidge A. Triple-Phase MDCT of Hepatocellular Carcinoma. Am J Roentgenol. 2004;182(3):643–9.
10. Marrero JA, Ahn J, Rajender Reddy K, Americal College of Gastroenterology. ACG clinical guideline: the diagnosis and management of focal liver lesions. Am J Gastroenterol. 2014;109(9):1328–47.
11. Jiang HY, Chen J, Xia CC, Cao LK, Duan T, Song B. Noninvasive imaging of hepatocellular carcinoma: From diagnosis to prognosis. World J Gastroenterol. 2018;24(22):2348–62.
12. Fukushima K, Kanemoto H, Ohno K, Takahashi M, Nakashima K, Fujino Y, et al. Ct characteristics of primary hepatic mass lesions in dogs. Vet Radiol Ultrasound. 2012;53(3):252–7.
13. Kutara K, Seki M, Ishikawa C, Sakai M, Kagawa Y, Iida G, et al. Triple-phase helical computed tomography in dogs with hepatic masses. Vet Radiol Ultrasound. 2014;55(1):7–15.
14. Jones ID, Lamb CR, Drees R, Priestnall SL, Mantis P. Associations between dual-phase

computed tomography features and histopathologic diagnoses in 52 dogs with hepatic or splenic masses. Vet Radiol Ultrasound. 2016;57(2):144–53.

15. Griebie ER, David FH, Ober CP, Feeney DA, Anderson KL, Wuenschmann A, et al. Evaluation of canine hepatic masses by use of triphasic computed tomography and B-mode, color flow, power, and pulsed-wave Doppler ultrasonography and correlation with histopathologic classification. Am J Vet Res. 2017;78(11):1273–83.
16. Leela Arporn R, Ohta H, Shimbo G, Hanazono K, Osuga T, Morishita K, et al. Computed tomographic features for differentiating benign from malignant liver lesions in dogs. J Vet Med Sci. 2019;81(12):1697–704.
17. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards Clinical Application of Image Mining: A Systematic Review on Artificial Intelligence and Radiomics. Eur J Nucl Med Mol Imaging. 2019;
18. Burti S, Longhin Osti V, Zotti A, Banzato T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. Vet J [Internet]. 2020;262:105505. Available from: https://doi.org/10.1016/j.tvjl.2020.105505
19. Banzato T, Bonsembiante F, Aresu L, Gelain ME, Burti S, Zotti A. Use of transfer learning to detect diffuse degenerative hepatic diseases from ultrasound images in dogs: A methodological study. Vet J [Internet]. 2018;233:35–40. Available from: http://dx.doi.org/10.1016/j.tvjl.2017.12.026
20. Banzato T, Bernardini M, Cherubini GB, Zotti A. A methodological approach for deep learning to distinguish between meningiomas and gliomas on canine MR-images. BMC Vet Res. 2018;14(1):317.
21. Banzato T, Cherubini GB, Atzori M, Zotti A. Development of a deep convolutional neural network to predict grading of canine meningiomas from magnetic resonance images. Vet J [Internet]. 2018;235:90–2. Available from: https://doi.org/10.1016/j.tvjl.2018.04.001
22. Banzato T, Zovi G, Milani C. Estimation of fetal lung development using quantitative analysis of ultrasonographic images in normal canine pregnancy. Theriogenology. 2017;96:158–63.
23. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: Current applications and future directions. PLoS Med. 2018;15(11):2–5.
24. Carlson SK, Johnson CD, Bender CE, Welch TJ. CT of focal nodular hyperplasia of the liver. Am J Roentgenol. 2000;174(3):705–12.
25. Sápi J, Kovács L, Drexler DA, Kocsis P, Gajári D, Sápi Z. Tumor volume estimation and quasi-continuous administration for most effective bevacizumab therapy. PLoS One. 2015;10(11):1–

20.

26. Therneau TM, Atkinson EJ. An introduction to recursive partitioning using the RPART routines. Encycl Med Decis Mak. 2019;1–60.

27. Fukushima K, Fujiwara R, Yamamoto K, Kanemoto H, Ohno K, Tsuboi M, et al. Characterization of triple-phase computed tomography in dogs with pancreatic insulinoma. J Vet Med Sci. 2016;77(12):1549–53.

28. Leela Arporn R, Ohta H, Shimbo G, Hanazono K, Osuga T, Morishita K, et al. Computed tomographic features for differentiating benign from malignant liver lesions in dogs. J Vet Med Sci. 2019;81(12):1697–704.

29. Rothuizen J, Twedt DC. Liver Biopsy Techniques. Vet Clin North Am Small Anim Pract [Internet]. 2009;39(3):469–80. Available from: http://www.sciencedirect.com/science/article/pii/S019556160900028X

30. McDevitt HL, Mayhew PD, Giuffrida MA, Brown DC, Culp WTN, Runge JJ. Short-term clinical outcome of laparoscopic liver biopsy in dogs: 106 cases (2003-2013). J Am Vet Med Assoc. 2016;248(1):83–90.

31. Webster CRL. Hemostatic Disorders Associated with Hepatobiliary Disease. Vet Clin North Am - Small Anim Pract [Internet]. 2017;47(3):601–15. Available from: http://dx.doi.org/10.1016/j.cvsm.2016.11.009

32. Lidbury JA. Getting the Most Out of Liver Biopsy. Vet Clin North Am - Small Anim Pract. 2017;47(3):569–83.

33. Bahr KL, Sharkey LC, Murakami T, Feeney DA. Accuracy of US-guided FNA of focal liver lesions in dogs: 140 cases (2005-2008). J Am Anim Hosp Assoc. 2013;49(3):190–6.

34. Wang KY, Panciera DL, Al-rukibat RK, Radi ZA. Accuracy of ultrasound-guided fine-needle aspiration of the liver and cytologic findings in dogs and cats: 97 cases (1990–2000). 2004;224(1):1–4.

35. Spangler WL, Culbertson MR. Prevalence, type, and importance of splenic diseases in dogs: 1,480 cases (1985-1989). J Am Vet Med Assoc. 1992;15(200(6)):829–34.

36. Kutara K, Seki M, Ishigaki K, Teshima K, Ishikawa C, Kagawa Y, et al. Triple-phase helical computed tomography in dogs with solid splenic masses. J Vet Med Sci. 2017;79(11):1870–7.

37. Meuten DJ. Tumors in Domestic Animals. Tumors in Domestic Animals. 2017.

38. Cordella A, Caldin M, Bertolini G. Splenic extramedullary hematopoiesis in dogs is frequently detected on multiphase multidetector-row CT as hypervascular nodules. Vet Radiol

Ultrasound. 2020;61(5):512–8.

39. Mattoon JSNTG. SMALL ANIMAL DIAGNOSTIC ULTRASOUND. 2015.

40. Fife WD, Samii VF, Drost T, Mattoon JS, Hoshaw-Woodard S. Comparison between malignant and nonmalignant splenic masses in dogs using contrast-enhanced computed tomography. Vet Radiol Ultrasound. 2004;45(4):289–97.

41. Lee M, Park J, Choi H, Lee H, Jeong SM. Presurgical assessment of splenic tumors in dogs: a retrospective study of 57 cases (2012–2017). J Vet Sci. 2018;19(6):827.

42. Banzato T, Fiore E, Morgante M, Manuali E, Zotti A, Manu- E, et al. Texture analysis of B-mode ultrasound images to stage hepatic lipidosis in the dairy cow: A methodological study. Res Vet Sci. 2016 Oct;108:71–5.

43. Banzato T, Bernardini M, Cherubini GB, Zotti A. Texture analysis of magnetic resonance images to predict histologic grade of meningiomas in dogs. Am J Vet Res. 2017 Oct;78(10):1156–62.

44. Li S, Wang Z, Visser LC, Wisner ER, Cheng H. Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. Vet Radiol Ultrasound. 2020;61(6):611–8.

45. McEvoy FJ, Proschowsky HF, Müller A V., Moorman L, Bender-Koch J, Svalastoga EL, et al. Deep transfer learning can be used for the detection of hip joints in pelvis radiographs and the classification of their hip dysplasia status. Vet Radiol Ultrasound. 2021;62(4):387–93.

46. Banzato T, Wodzinski M, Burti S, Osti VL, Rossoni V, Atzori M, et al. Automatic classification of canine thoracic radiographs using deep learning. Sci Rep [Internet]. 2021;11(1):1–8. Available from: https://doi.org/10.1038/s41598-021-83515-3

47. Banzato T, Causin F, Della Puppa A, Cester G, Mazzai L, Zotti A. Accuracy of deep learning to differentiate the histopathological grading of meningiomas on MR images: A preliminary study. J Magn Reson Imaging. 2019;50(4):1152–9.

48. Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. Front Vet Sci. 2021;8(March):1–10.

49. Mangano C, Macrì F, Di Pietro S, Pugliese M, Santoro S, Iannelli NM, et al. Use of contrast-enhanced ultrasound for assessment of nodular lymphoid hyperplasia (NLH) in canine spleen. BMC Vet Res. 2019;15(1):1–9.

50. Rossi F, Leone VF, Vignoli M, Laddaga E, Terragni R. Use of contrast-enhanced ultrasound for characterization of focal splenic lesions. Vet Radiol Ultrasound. 2008;49(2):154–64.

51. Barat M, Hoeffel C, Aissaoui M, Dohan A, Oudjit A, Dautry R, et al. Focal splenic lesions: Imaging spectrum of diseases on CT, MRI and PET/CT. Diagn Interv Imaging. 2021;102(9):501–13.
52. Clifford CA, Pretorius ES, Weisse C, Sorenmo KU, Drobatz KJ, Siegelman ES, et al. Magnetic Resonance Imaging of Focal Splenic and Hepatic Lesions in the Dog. J Vet Intern Med. 2004;18(3):330–8.
53. Tecilla M, Gambini M, Forlani A, Caniatti M, Ghisleni G, Roccabianca P. Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. PLoS One. 2019;14(11):1–15.

CHAPTER THREE

Computed tomography features for differentiating malignant and benign focal liver lesions in dogs: a meta-analysis.

The chapter was adapted from:

Burti S, Zotti A, Contiero B, Banzato T. *Computed tomography features for differentiating malignant and benign focal liver lesions in dogs: A meta-analysis*. Vet J. 2021 Dec;278:105773. doi: 10.1016/j.tvjl.2021.105773. Epub 2021 Nov 3. PMID: 34742915.

ABSTRACT

Computed tomography (CT) is often performed to complement ultrasound following detection of focal liver lesions (FLL). There is no consensus in the literature regarding the CT features that might be helpful in the distinction between benign and malignant FLL. The aim of this meta-analysis was to identify, based on the available literature, the qualitative and quantitative CT features able to distinguish between benign and malignant FLL. Studies on the diagnostic accuracy of CT in characterising FLL were searched in MEDLINE, Web of Science, and Scopus databases. Pooled sensitivity, pooled specificity, diagnostic odds ratio (DOR), receiver operator curve (ROC) area, were calculated for qualitative features. DORs were used to determine which qualitative features were most informative to detect malignancy; quantitative features were selected/identified based on standardised mean difference (SMD).

Well-defined margins, presence of a capsule, abnormal lymph nodes, and heterogeneity in the arterial, portal and delayed phase were classified as informative qualitative CT features. The pooled sensitivity ranged from 0.630 (abnormal lymph nodes) to 0.786 (well-defined margins), while pooled specificity ranged from 0.643 (well-defined margins) to 0.816 (heterogeneous in delayed phase). Maximum dimensions, ellipsoid volume, attenuation of the liver in the pre-contrast phase, and attenuation of the liver in the arterial, portal, and delayed phase were found to be informative quantitative CT features. Larger maximum dimensions and volume (positive SMD), and lower attenuation values (negative SMD) were more associated with malignancy. This meta-analysis provides the evidence base for the interpreting CT imaging in the characterization of FLL.

INTRODUCTION

Computed tomography (CT) is widely used for the diagnosis and monitoring of many diseases in dogs. CT is also frequently used as a stand-alone diagnostic imaging technique for neoplasm staging. Masses or nodules may also be found incidentally during CT imaging performed to investigate diseases elsewhere in the body (Burti et al., 2021). Regardless of reason for scanning, focal liver lesions (FLL) are common findings on CT scans of dogs, especially when older animals are investigated (Jones et al., 2016). FLL may also be initially identified using other diagnostic imaging techniques, such as ultrasonography, and then, if appropriate, better characterised by means of CT (Marolf, 2017). CT offers the ability to evaluate the liver and any lesions in three dimensions and provides superior visualisation of lesional vascularisation in comparison to ultrasound. Furthermore, some types of lesions (e.g. vacuolar degeneration) are detectable only through enhanced CT techniques. While the CT features of the different histotypes of FLL have been widely investigated in human medicine, and, especially in the case of hepatocellular carcinoma (HCC; Shah et al., 2014), are well known and characterised (Ariff et al., 2009; Chou et al., 2015); to date, the CT features of FLL in dogs have been infrequently described. In addition, although some features (e.g., enhancement in the delayed phase, lesion dimensions) have been reported as useful in differentiating between benign and malignant FLL in dogs (Griebie et al., 2017; Burti et al., 2021), other authors report that no CT features were useful to aid this differentiation (Stehlík et al., 2020). The usefulness of CT in discriminating between benign and malignant FLL is still to be fully determined.

The aim of this meta-analysis was to identify qualitative and quantitative CT features useful in distinguishing between benign and malignant FLL and to summarise their diagnostic accuracy.

MATERIALS AND METHODS

This meta-analysis was performed in accordance with the best practices in diagnostic test accuracy systematic reviews (PRISMA-DATA; Moher et al., 2009; DerSimonian and Laird, 2015).

Search Strategy

A literature search was conducted based on the PICOS (population, intervention, comparator, outcome, study design) approach (Methley et al., 2014). The target population was dogs with FLL evident on CT scans that had histopathological analysis of the FLL performed. Intervention was identification of CT features. Outcome was the diagnostic accuracy on benign vs. malignant characterisation of the FLL. FLL were defined as liver nodule(s) or mass(es) of any dimensional value, that were identified in the CT scans due to their different attenuation when compared to the surrounding liver parenchyma. Study design was controlled or comparative, randomised or non-randomised experimental studies, or prospective or retrospective observational studies. A search of the MEDLINE, Web of Science, and Scopus databases from January 2000 to Jan 2021 was performed. To maximise inclusion of articles for review, we opted to perform a search based on generic terms. Keywords used for searching were (computed tomography OR "CT" AND liver OR hepatic AND dog OR canine). The literature search was restricted to articles written in the English language.

Screening of studies

Screening of the studies was performed by two authors. The studies were first screened at both title and abstract level. Reviews were excluded. Thereafter, the identified articles were selected at a full-text level and only those that entirely met the PICOS criteria were included.

Eligibility criteria

The inclusion criteria for the studies were: (1) CT evaluation of FLL in dogs; (2) evaluation of qualitative and quantitative CT features; (3) cytological and/or histological diagnosis of the lesions.

Reference standard

The reference standard were lesions with cytological and/or histological-confirmed diagnosis. Studies that used either histopathology and/or cytology as reference standards were included. Reference standards that fulfilled the above criteria were considered at low risk of bias.

Data extraction

For each included study, the following characteristics were recorded: type of study (i.e. prospective or retrospective), country in which the study was performed, time period over which data were collected, CT scanning method (i.e. dual-phase or triple-phase), number of dogs or FLL, age, body weight, and FLL size. In addition, the cytopathological/histopathological categories used, along with the number of FLL within each category, were also recorded. The cytopathological/histopathological categories used were condensed into 'benign' (which included nodular hyperplasia [NH]) or 'malignant' (which included HCC).

Risk of bias assessment

The risk of bias was assessed using the QUADAS-2 tool (Whiting et al., 2016). No modification to this tool was necessary for the specific search question. The assessment was completed independently by two authors. Discrepancies were resolved with the aid of a co-author of this study.

Outcomes

The primary outcome was the sensitivity and specificity of individual CT features in the identification of malignant FLL in dogs. Data were analysed on a per-lesion level, as both benign and malignant lesions can be present in the same individual at the same time. Secondary outcomes were CT characteristics, study population factors, study design, and risk of bias.

Data analysis

For dichotomous predictor variables, diagnostic test measures were calculated using 2 x 2 contingency tables based on the number of FLL displaying each CT feature as reported in the results of the articles included in the meta-analysis. Sensitivity, specificity, and accuracy, along with their 95% confidence intervals (CI), were calculated for each individual study. The area under the receiver operating characteristic (ROC) curves (AUC) was also estimated, and this information was used for meta-analysis of the qualitative features. The outcome of the meta-analysis was a pooled estimate of sensitivity, specificity, diagnostic odds ratio (DOR), and AUC. Based on these aggregated results, it was possible to determine whether the presence of a certain CT feature was an accurate predictor for malignancy. DOR could range from 0 to infinity, with higher values indicating better discriminatory test performance. As value of 1 implies that the test had no discriminative power, any CT feature where the 95% CI of its DOR did not span 1 was considered as informative.

For the continuous predictors, the descriptive statistics (as mean and standard deviation (SD); or median and range) were reported for the two groups (malignant vs. benign) as reported in their respective individual studies. For the meta-analysis of the continuous measures, comparison of the means between malignant and benign cases was performed using the standardised mean difference (SMD). For the studies reporting only the range the SD was estimated by dividing the range by 4. Based on the pooled result, any CT features where the 95% CI of its SMD did not span 0 were considered statistically significant at the 5% level ($P < 0.05$). Cohen's rule of thumb for interpreting the SMD statistic is to consider the absolute value: a value of 0.2 indicates a small effect, a value of 0.5 indicates a medium effect and a value of 0.8 or higher indicates a large effect.

For both qualitative and quantitative CT features, study heterogeneity was assessed to determine whether a fixed or random effects model had to be used for the meta-analysis. The agreement or disagreement between the studies was examined using different measures of heterogeneity: Cochran's Q and I^2 statistics (Higgins, 2003). A $Q > 0.1$ and an $I^2 > 0.5$ were considered as indicative for heterogeneity. When heterogeneity was present, the random effects model was used.

Forest plot graphs were used to show the meta-analysis results for every study, along with the 95% CI and the numerical estimate of the overall effect of interest (global DOR or SMD, sensitivity, and specificity). In the graphs, the length of the horizontal lines represents the confidence intervals of the studies, the dimensions of the boxes represent the weights assigned to each of them. These weights depended on sample size and on the model adopted (fixed or random effects). All the analyses were conducted using 'mada' (Doebler, 2020) and 'meta' (Schwarzer, 2021) packages (Shim et al., 2019) of R (version: 2020)[1].

RESULTS

Study search

The search in the MEDLINE, Web of Science, and Scopus databases retrieved 435 potentially relevant studies. Duplicates were removed. All the articles not matching the inclusion criteria (case reports, reviews, letters, abstracts, recommendations, guidelines) were excluded (*n* = 427). Eight articles in total matched the inclusion criteria. The study selection process is reported in Fig.1.

---

[1] See: R: A Language and Environment for Statistical Computing https://www.R-project.org (Accessed 25 October 2021).

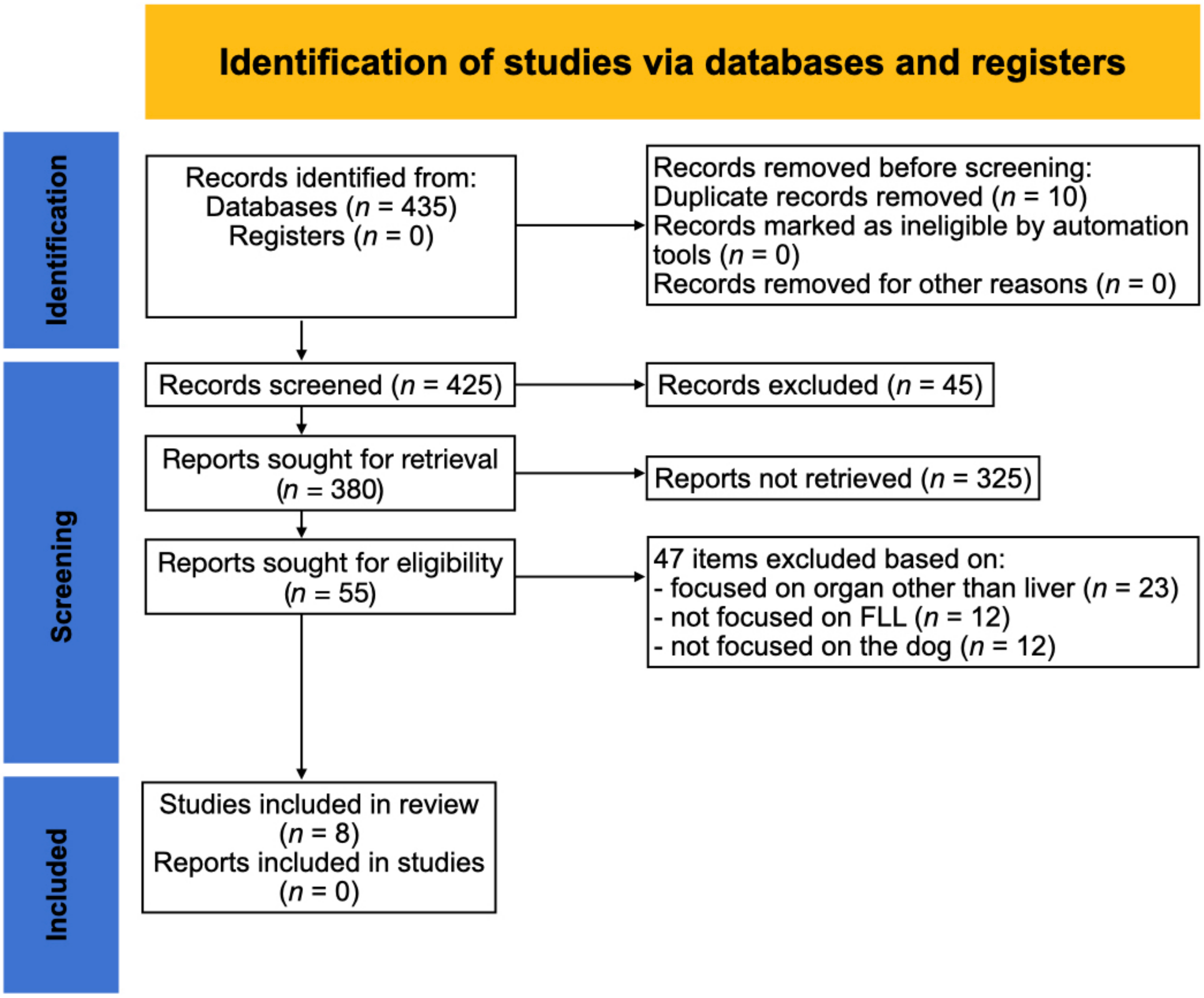


Figure 1 Flow chart showing the literature search process

Study characteristics

A total of 404 dogs were included, and the CT features of 419 FLL, along with their histopathological or cytological diagnosis, were reported. In Taniura et al. (2009) only FLL with a diagnosis of nodular hyperplasia (NH) or HCC were included. In Fukushima et al. (2012) only FLL with a diagnosis of HCC, NH, or another benign process were included. No a-priori selection of the FLL was made in the remaining studies. The characteristics of the studies are summarised in Table 1.

Table 1 Characteristics of the studies included in the analysis.

	Study design/Country	Time period	Scanning method	Dogs/FLL (n)	Age range (years)	Body weight (kg)	Diagnosis (n)	Lesion size (cm)
Stehlik et al., 2021	Retrospective / Czech Republic and Italy	2016-2019	Triple-phase CT	31	5-16 years	23.3 ± 10.8 ^{a}	Benign (17) Malignant (14)	
Burti et al., 2021	Retrospective / Italy	2015-2020	Dual-phase CT	69	4-16 years	Not given	Benign (37) Malignant (32)	5.37 ± 4.36 ^{a}
Leela-Arporn et al., 20	Prospective / Japan	2016-2019	Triple-phase CT	57 / 70	8-13 years	9.2 ± 6.6 ^{a}	Benign (18) Malignant (52)	4.85 ± 2.25 $^{a, c}$
Griebie et al., 2017	Retrospective / USA	2014-2016	Triple-phase CT	44 / 46	4-13 years	Not given	Benign (16) Malignant (30)	7.135 ^{c}
Jones et al., 2016	Retrospective/ UK	2008-2014	Dual-phase CT	24	3-15 years	24 (4-54) ^{b}	Benign (10) Malignant (14)	6.0 (1.5-6.95) ^{b}
Kutara et al., 2014	Prospective / Japan	Not given	Triple-phase CT	70	7-16 years	14.8 ± 9.1 ^{a}	Benign (14) Malignant (56)	6.35 ± 3.32 ^{a}
Fukushima et al., 201	Retrospective / Japan	2005-2010	Triple-phase CT	33	0.25-15 years	7.2 (1.5-38.5) ^{b}	Benign (19) Malignant (14)	7.77 (2.8-16.5) ^{b,}
Taniura et al., 2009	Retrospective / Japan	2004-2007	Triple-phase CT	76	Not given	Not given	Benign (40) Malignant (36)	4.76 ± 2.06 $^{a, c}$

CT, computed tomography; FLL, focal liver lesions

[a] Mean (± standard deviation)

[b] Median (range)

[c] Pooled results from available data

Quality of the studies

The risk of bias was evaluated as high for the selection of cases in the studies by Taniura et al. (2009) and Kutara et al. (2014), as only dogs diagnosed with certain pathologies (NH and HCC in Taniura et al., 2009; HCC, NH and metastatic tumours in Kutara et al., 2014) were included. The applicability of the results presented, as well as the applicability of the index test, were consequentially considered as heavily biased. For the remaining studies the risk of bias was classified as low for: case selection, Index test, and flow and timing. For all the studies, the blinding of pathologists and radiologists to the results of other tests was not mentioned and, therefore, classified as unknown. Lastly, a power analysis was not conducted in any of the included studies. The study quality results are summarised in Table 2.

Table 2 Risk of bias and applicability concerns assessment for each of the studies included in the analysis

	Risk of bias				Applicability concerns		
	Case selection	**Index test**	**Reference standard**	**Flow and timing**	**Case selection**	**Index test**	**Reference standard**
Taniura et al., 2009	High	Unknown	Unknown	Low	High	Unknown	Unknown
Fukushima et al., 2012	Low	Low	Unknown	Low	Low	Low	Low
Kutara et al., 2014	High	Unknown	Unknown	Low	High	Unknown	Low
Jones et al., 2016	Low	Low	Unknown	Low	Low	Low	Low
Griebie et al., 2017	Low	Low	Unknown	Low	Low	Low	Low
Leela-Arporn et al., 2019	Low	Low	Unknown	Low	Low	Low	Low
Stehlik et al., 2021	Low	Low	Unknown	Low	Low	Low	Low
Burti et al., 2021	Low	Low	Unknown	Low	Low	Low	Low

Categorisation of the CT features

Twenty-eight overlapping CT features were evaluated in the different studies. Sixteen qualitative features (well-defined margins, irregular surface, presence of a capsule, abnormal lymph nodes, marginal enhancement pattern, diffuse enhancement pattern, heterogeneous pattern in arterial phase, heterogeneous pattern in portal phase, heterogeneous pattern in delayed phase, hypoattenuation, and hypoenhancement, enhancement in the portal phase, and enhancement in the delayed phase; these latter two features were further divided into hypo, iso, and hyper) and 10 quantitative continuous features (maximum dimension, ellipsoid volume, attenuation of normal liver in the pre-contrast phase, attenuation of normal liver in arterial phase, attenuation of normal liver in portal phase, attenuation of normal liver in delayed phase, attenuation of pre-contrast FLL, attenuation of FLL in arterial phase, attenuation of FLL in portal phase, attenuation of FLL in delayed phase) were considered because these were evaluated in at least two studies.

Diagnostic accuracy of the CT features in the individual studies

The sensitivity, specificity, accuracy, and AUC of the qualitative CT features as evaluated in each study are reported in Table 3. The mean and the standard deviation and/or median with overall range of the continuous variables as evaluated in the individual studies are reported in Table 4. Taniura et al. (2009) reported AUCs above 0.80 for most of the considered CT features, whereas Kutara et al. (2014) reported an AUC of 0.82 for heterogeneity in the portal phase. Accuracy index values higher than 80% were evident for well-defined margins in Leela-Arpon et al. (2019) and Taniura et al. (2009), and for presence of a capsule, hypoattenuation, and hypoenhancement in the portal phase and delayed phase in Taniura et al. (2009). Significant differences for the quantitative CT features were reported for the maximum lesion dimensions in Leela-Arpon et al. (2019) and Taniura et al. (2009), for attenuation of pre-contrast normal liver in Burti et al. (2021), and for ellipsoid volume, attenuation of post-contrast normal liver, and attenuation of normal liver in portal phase in Leela-Arpon et al. (2019).

Table 3 Calculated accuracy measures of malignant diagnosis considering the qualitative features (95% CI, 95% confidence interval; SE, standard error; AUC, area under the curve).

Qualitative feature and references	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Accuracy (95% CI)**	**AUC**	**SE**
Well-defined margins					
Burti et al., 2021	0.750 (0.566-0.885)	0.487 (0.319-0.656)	0.608 (0.483-0.724)	0.618	0.07
Leela-Arporn et al., 2019	0.933 (0.841-0.988)	0.933 (0.173-0.643)	0.800 (0.687-0.886)	0.666	0.07
Jones et al., 2016	0.500 (0.230-0.700)	0.600 (0.262-0.878)	0.542 (0.328-0.745)	0.55	0.12
Fukushima et al., 2012	0.857 (0.572-0.982)	0.211 (0.061-0.456)	0.485 (0.308-0.665)	0.534	0.1
Taniura et al., 2009	0.694 (0.519-0.837)	1 (0.912-1)	0.855 (0.756-0.926)	0.847	0.04
Irregular surface					
Burti et al, 2021	0.938 (0.792-0.992)	0.432 (0.271-0.605)	0.667 (0.543-0.776)	0.685	0.06
Leela-Arporn et al., 2019	0.346 (0.220-0.491)	0.940 (0.727-0.999)	0.500 (0.378-0.622)	0.645	0.07
Fukushima et al., 2012	0.571 (0.289-0.823)	0.211 (0.061-0.456)	0.364 (0.204-0.549)	0.527	0.1
Presence of a capsule					
Burti et al, 2021	0.500 (0.319-0.681)	0.811 (0.648-0.920)	0.667 (0.543-0.776)	0.655	0.07
Leela-Arporn et al., 2019	0.423 (0.287-0.568)	0.889 (0.653-0.986)	0.543 (0.419-0.663)	0.656	0.07
Fukushima et al., 2012	0.929 (0.661-0.998)	0.474 (0.245-0.711)	0.667 (0.482-0.820)	0.701	0.09
Taniura et al., 2009	0.694 (0.519-0.837)	1 (0.912-1)	0.855 (0.756-0.926)	0.847	0.05
Abnormal lymph nodes					
Burti et al., 2021	0.563 (0.377-0.736)	0.838 (0.680-0.938)	0.710 (0.588-0.813)	0.7	0.06
Jones et al., 2016	0.785 (0.492-0.953)	0.400 (0.122-0.738)	0.625 (0.406-0.812)	0.593	0.12
Marginal enhancement pattern					
Burti et al., 2021	0.156 (0.053-0.328)	0.811 (0.648-0.920)	0.507 (0.384-0.630)	0.484	0.07
Jones et al., 2016	0.143 (0.018-0.428)	0.800 (0.444-0.975)	0.417 (0.221-0.634)	0.471	0.12
Diffuse enhancement pattern					
Burti et al., 2021	0.813 (0.636-0.928)	0.189 (0.080-0.352)	0.478 (0.357-0.602)	0.501	0.07
Jones et al., 2016	0.857 (0.572-0.982)	0.200 (0.025-0.556)	0.583 (0.366-0.779)	0.529	0.12
Hypoattenuation					
Burti et al., 2021	0.875 (0.710-0.965)	0.243 (0.118-0.412)	0.536 (0.412-0.657)	0.559	0.06
Taniura et al., 2009	0.833 (0.672-0.936)	0.850 (0.702-0.943)	0.842 (0.740-0.916)	0.842	0.04
Heterogeneous in arterial phase					
Leela-Arporn et al., 2019	0.769 (0.632-0.875)	0.556 (0.308-0.785)	0.714 (0.594-0.816)	0.662	0.1
Kutara et al., 2014	0.714 (0.578-0.827)	0.571 (0.289-0.823)	0.686 (0.564-0.792)	0.643	0.11
Heterogeneous in portal phase					
Burti et al., 2021	0.813 (0.636-0.928)	0.514 (0.344-0.681)	0.652 (0.528-0.763)	0.663	0.07
Leela-Arporn et al., 2019	0.769 (0.632-0.875)	0.722 (0.465-0.903)	0.757 (0.640-0.852)	0.746	0.06
Jones et al., 2016	0.429 (0.177-0.711)	0.400 (0.122-0.738)	0.417 (0.221-0.634)	0.714	0.1
Kutara et al., 2014	0.714 (0.578-0.827)	0.929 (0.661-0.998)	0.757 (0.640-0.852)	0.821	0.05

Heterogeneous in delayed phase					
Leela-Arporn et al., 2019	0.750 (0.611-0.860)	0.722 (0.465-0.903)	0.743 (0.624-0.840)	0.736	0.06
Kutara et al., 2014	0.554 (0.415-0.687)	0.929 (0.661-0.998)	0.629 (0.505-0.741)	0.741	0.07
Hyponhancement in portal phase					
Stehlik et al., 2021	0.429 (0.177-0.711)	0.412 (0.184-0.671)	0.419 (0.246-0.609)	0.42	0.1
Leela-Arporn et al., 2019	0.096 (0.032-0.210)	0.833 (0.586-0.964)	0.286 (0.184-0.406)	0.465	0.08
Jones et al., 2016	0.143 (0.018-0.428)	0.800 (0.444-0.975)	0.417 (0.221-0.634)	0.471	0.12
Kutara et al., 2014	0.571 (0.432-0.703)	1 (0.768-1)	0.657 (0.534-0.766)	0.786	0.06
Fukushima et al., 2012	0.857 (0.572-0.982)	0.421 (0.203-0.665)	0.606 (0.421-0.771)	0.639	0.1
Taniura et al., 2009	0.917 (0.775-0.983)	0.975 (0.868-0.999)	0.947 (0.871-0.986)	0.946	0.03
Isoenhancement in portal phase					
Stehlik et al., 2021	0.571 (0.289-0.823)	0.706 (0.440-0.897)	0.645 (0.454-0.808)	0.639	0.1
Leela-Arporn et al., 2019	0.154 (0.069-0.281)	0.778 (0.524-0.936)	0.314 (0.209-0.436)	0.466	0.08
Kutara et al., 2014	0.268 (0.158-0.403)	0.643 (0.351-0.872)	0.343 (0.234-0.466)	0.455	0.09
Fukushima et al., 2012	0.143 (0.018-0.428)	0.842 (0.604-0.966)	0.546 (0.364-0.719)	0.492	0.1
Taniura et al., 2009	0.056 (0.068-0.187)	0.075 (0.016-0.204)	0.066 (0.022-0.147)	0.065	0.03
Hyperenhancement in portal phase					
Stehlik et al., 2021	0 (0-0.232)	0.882 (0.636-0.985)	0.484 (0.301-0.669)	0.441	0.1
Leela-Arporn et al., 2019	0.750 (0.611-0.860)	0.389 (0.173-0.643)	0.657 (0.534-0.767)	0.569	0.08
Jones et al., 2016	0.500 (0.230-0.770)	0.900 (0.555-0.998)	0.667 (0.447-0.844)	0.7	0.11
Kutara et al., 2014	0.161 (0.076-0.283)	0.357 (0.128-0.649)	0.200 (0.114-0.313)	0.259	0.08
Fukushima et al., 2012	0 (0-0.232)	0.737 (0.488-0.909)	0.424 (0.255-0.608)	0.368	0.1
Taniura et al., 2009	0.028 (0-0.145)	1 (0.912-1)	0.539 (0.421-0.654)	0.514	0.07
Hyponhancement in delayed phase					
Stehlik et al., 2021	0.357 (0.128-0.649)	0.529 (0.278-0.770)	0.452 (0.273-0.640)	0.443	0.11
Burti et al., 2021	0.906 (0.750-0.980)	0.162 (0.062-0.320)	0.507 (0.384-0.630)	0.534	0.07
Leela-Arporn et al., 2019	0.039 (0.005-0.132)	0.833 (0.586-0.964)	0.243 (0.148-0.360)	0.436	0.08
Kutara et al., 2014	0.482 (0.347-0.620)	1 (0.768-1)	0.586 (0.462-0.702)	0.741	0.07
Fukushima et al., 2012	0.929 (0.661-0.998)	0.421 (0.202-0.665)	0.636 (0.451-0.796)	0.675	0.1
Taniura et al., 2009	0.944 (0.813-0.993)	1 (0.912-1)	0.974 (0.908-0.997)	0.972	0.02
Isoenhancement in delayed phase					
Stehlik et al., 2021	0.571 (0.289-0.823)	0.471 (0.229-0.722)	0.516 (0.331-0.699)	0.521	0.11
Burti et al., 2021	0 (0-0.109)	0.973 (0.858-0.999)	0.522 (0.384-0.630)	0.486	0.07
Leela-Arporn et al., 2019	0.308 (0.187-0.451)	0.556 (0.308-0.785)	0.371 (0.259-0.495)	0.432	0.08
Kutara et al., 2014	0.393 (0.265-0.533)	0.286 (0.084-0.581)	0.371 (0.259-0.495)	0.339	0.09
Fukushima et al., 2012	0.071 (0.002-0.339)	0.632 (0.384-0.837)	0.394 (0.229-0.579)	0.352	0.1
Taniura et al., 2009	0.056 (0.007-0.187)	0 (0-0.088)	0.026 (0.003-0.092)	0.028	0.02
Hyperenhancement in delayed phase					
Stehlik et al., 2021	0.071 (0.002-0.339)	1 (0.805-1)	0.581 (0.391-0.754)	0.536	0.11
Burti et al., 2021	0.062 (0.008-0.208)	0.865 (0.712-0.955)	0.493 (0.370-0.616)	0.464	0.07
Leela-Arporn et al., 2019	0.654 (0.509-0.780)	0.611 (0.358-0.827)	0.643 (0.519-0.754)	0.632	0.07
Kutara et al., 2014	0.125 (0.052-0.241)	0.714 (0.419-0.916)	0.243 (0.148-0.360)	0.42	0.09
Fukushima et al., 2012	0 (0-0.232)	0.947 (0.740-0.999)	0.546 (0.363-0.719)	0.474	0.1

Table 4 Descriptive values for malignant and benign lesions recorded as quantitative continuous features with reported P-values.

Qualitative feature and references	**Malignant**	**Benign**	**Reported *P*-value**
Maximum dimension (cm)			
Burti et al., 2021	4.3 (0.5-16.3)	3.5 (0.5-18.1)	0.06
Leela-Arporn et al, 2019	6.6 (±3.1)	3.1 (±1.4)	<0.0001
Jones et al, 2016	6.1 (1.5-69.5)	6.0 (3.2-34.5)	0.95
Taniura et al., 2009	6.8 (±3.5)	2.7 (±0.6)	<0.05
Ellipsoid volume (cm^3)			
Burti et al., 2021	20.8 (0.02-1576)	11.5 (0.04-1995)	0.2
Leela-Arporn et al., 2019	195.0 (±228.6)	21.3 (±24.5)	<0.0001
Attenuation of pre-contrast normal liver (HU)			
Stehlik et al., 2021	63.1 (±9.0)	60.9 (±8.7)	n.s.
Burti et al., 2021	58.9 (46.2-82.1)	66.1 (51.8-85.5)	< 0.01
Jones et al., 2016	60.0 (53.0-75.0)	71.0 (49.0-79.0)	0.13
Taniura et al., 2009	58.1 (±11.2)	62.7 (±8.1)	Not reported
Attenuation of pre-contrast lesion (HU)			
Burti et al., 2021	40.3 (10.5-67.3)	43.7 (9.6-70.1)	0.66
Jones et al., 2016	46.0 (32.0-68.0)	60.0 (26.0-69.0)	0.32
Taniura et al., 2009	44.2 (±11.6)	61.1 (±10.1)	Not reported
Attenuation of arterial phase – normal liver (HU)			
Stehlik et al., 2021	67.9 (±7.1)	69.9 (±16.1)	n.s.
Leela-Arporn et al., 2019	117.2 (±18.4)	121.2 (±16.1)	0.39
Taniura et al., 2009	106.1 (±26.4)	111.9 (±24.8)	Not reported
Attenuation of arterial phase – lesion (HU)			
Stehlik et al., 2021	52.8 (±17.2)	65.0 (±28.8)	n.s.
Leela-Arporn et al., 2019	112.0 (±49.6)	134.7 (±72.3)	0.15
Taniura et al., 2009	91.1 (±35.2)	116.1 (±44.4)	Not reported
Attenuation of portal phase – normal liver (HU)			
Stehlik et al., 2021	113.8 (±11.4)	114.6 (±18.7)	n.s.
Leela-Arporn et al., 2019	156.0 (±25.5)	172.1 (±26.3)	0.03
Taniura et al., 2009	136.6 (±23.7)	126.0 (±21.8)	Not reported
Attenuation of portal phase – lesion (HU)			
Stehlik et al., 2021	85.6 (±30.6)	90.5 (±42.4)	n.s.
Leela-Arporn et al., 2019	118.4 (±44.9)	148.4 (±61.1)	0.07
Taniura et al., 2009	109.8 (±38.9)	130.8 (±24.3)	Not reported

Attenuation of delayed phase – normal liver (HU)			
Stehlik et al., 2021	107.1 (±13.2)	112.8 (±7.1)	n.s.
Leela-Arporn et al., 2019	125.3 (±16.8)	131.1 (±15.4)	0.19
Taniura et al., 2009	127.7 (±22.1)	122.9 (±17.7)	Not reported
Attenuation of delayed phase – lesion (HU)			
Stehlik et al., 2021	87.0 (±30.2)	83.2 (±29.3)	n.s.
Burti et al, 2021	83.2 (62.4-121.2)	83.9 (62.6-121.3)	0.13
Leela-Arporn et al., 2019	105.2 (± 28.1)	119.7 (± 37.5)	0.15
Jones et al., 2016	77.0 (42.0-120.0)	99.0 (25.0-121.0)	0.62
Taniura et al., 2009	104.9 (±34.6)	123.9 (±17.5)	Not reported

n.s., not statistically significant ($p > 0.05$); HU, Hounsfield unit; CE, contrast-enhanced

[a] Data are reported as mean (± standard deviation) or as median (range)

Overall diagnostic accuracy

Results of the meta-analysis for qualitative CT features reported in at least two studies are reported in Table 5. The pooled sensitivity, specificity, DOR, and AUC were calculated using either a fixed or random effects model, depending on the outcome of the heterogeneity test. The results of the heterogeneity tests for each variable considered are reported as Supplementary Material. Six of the 16 qualitative features included in the meta-analysis were considered as informative. These were: well-defined margins, presence of a capsule, abnormal lymph nodes, and heterogeneity in the arterial, portal, and delayed phase. Of these, presence of a capsule and hyperenhancement in the delayed phase showed the highest specificity, with 0.884 (95% CI 0.537-0.980) and 0.864 (95% CI 0.681-0.950) respectively. These results were considered quite robust, as they were based on four and five studies respectively. Two variables showed a pooled sensitivity above 0.8, these were: hypoattenuation (0.853, 95% CI 0.748-0.919) and diffuse enhancement pattern (0.826, 95% CI 0.689-0.911). However, these results were based on only two studies each, with hypoattenuation evaluated in Burti et al. (2021) and Taniura et al. (2009), and diffuse enhancement pattern evaluated in Burti et al. (2021) and Jones et al. (2016). Heterogeneity in the portal phase showed the highest overall diagnostic accuracy, with an AUC of 0.751 and a DOR of 4.749. The forest plots of the pooled diagnostic accuracy measures (sensitivity, specificity, and DOR) for heterogeneity in the portal phase is reported in Fig.2. The remaining forest plots graphs are reported as Supplementary Material.

Table 5 Summary of meta-analysis for qualitative features: diagnostic accuracy of the predictors to identify the malignant cases. Overall sensitivity, specificity, and diagnostic odd's ratio (DOR) are reported with 95% confidence interval between parentheses. Overall area under the curve (AUC) and standard error (SE) are reported.

CT features	**Studies (n)**	**Sensitivity**	**Specificity**	**DOR**	**AUC**	**SE**
Well-defined margins	5	0.790 (0.622-0.890)	0.641 (0.220-0.921)	4.83 (1.43-16.36)	0.677	0.06
Irregular surface	3	0.680 (0.291-0.922)	0.580 (0.163-0.909)	3.18 (0.30-33.29)	0.535	0.07
Capsule presence	4	0.631 (0.420-0.809)	0.883 (0.537-0.980)	9.61 (2.75-33.55)	0.728	0.05
Abnormal lymph nodes	2	0.633 (0.483-0.760)	0.678 (0.331-0.901)	5.02 (1.94-12.99)	0.676	0.06
Marginal enhancement pattern	2	0.151 (0.070-0.290)	0.806 (0.670-0.910)	0.76 (0.26-2.25)	0.481	0.06
Diffuse enhancement pattern	2	0.834 (0.692-0.914)	0.194 (0.103-0.332)	1.11 (0.39-3.19)	0.508	0.06
Hypoattenuation	2	0.850 (0.750-0.919)	0.571 (0.153-0.913)	8.04 (0.67-96.26)	0.708	0.1
Heterogeneous in arterial phase	2	0.740 (0.650-0.815)	0.523 (0.392-0.721)	3.75 (1.64-8.57)	0.654	0.07
Heterogeneous in portal phase	4	0.730 (0.653-0.792)	0.660 (0.431-0.831)	4.75 (1.16-19.42)	0.751	0.03
Heterogeneous in delayed phase	2	0.652 (0.511-0.780)	0.820 (0.610-0.930)	9.34 (3.28-26.61)	0.738	0.04
Hypoenhancement in portal phase	6	0.510 (0.230-0.811)	0.831 (0.533-0.964)	4.22 (0.58-30.91)	0.64	0.08
Isoenhancement in portal phase	5	0.200 (0.091-0.373)	0.599 (0.281-0.853)	0.37 (0.05-2.51)	0.407	0.09
Hyperenhancement in portal phase	6	0.102 (0.011-0.503)	0.810 (0.480-0.950)	0.71 (0.14-3.61)	0.472	0.06
Hypoenhancement in delayed phase	6	0.663 (0.243-0.921)	0.832 (0.313-0.984)	4.83 (0.66-35.32)	0.646	0.08
Isoenhancement in delayed phase	6	0.151 (0.043-0.423)	0.442 (0.091-0.869)	0.20 (0.05-0.83)	0.345	0.07
Hyperenhancement in delayed phase	5	0.110 (0.020-0.390)	0.861 (0.683-0.950)	0.93 (0.30-2.89)	0.511	0.04

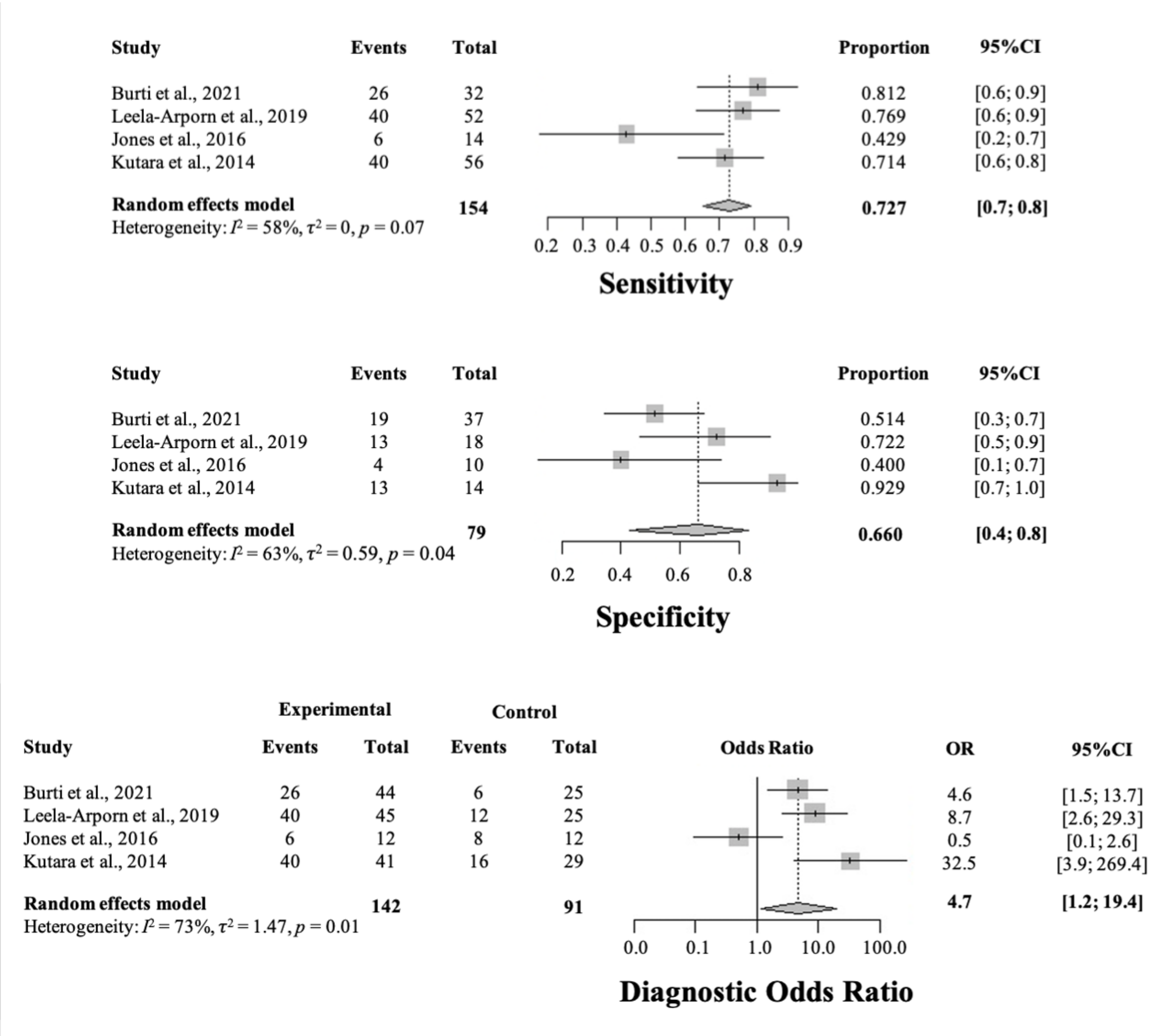


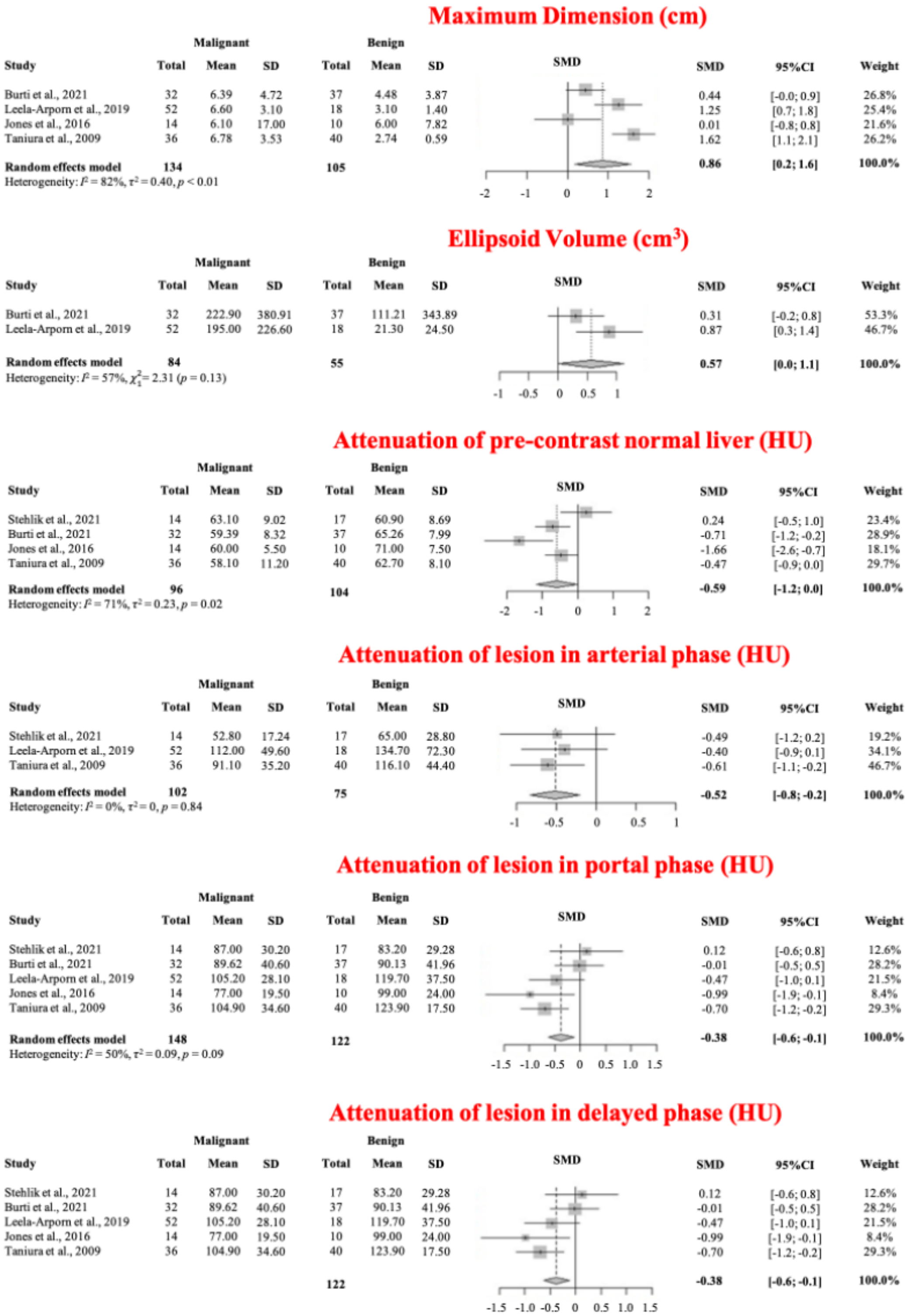
Figure 2 Forest plot of the pooled diagnostic accuracy measures (sensitivity, specificity, and diagnostic odds ratio [DOR]) for heterogeneity in the portal phase. The squares represent the proportion of malignant focal liver lesion (FLL) and the whiskers represent the 95% confidence interval (CI). The diamond represents the pooled effect. The location of the diamond represents the estimated effect size, and the width of the diamond reflects the precision of the estimate. Heterogeneity indexed (I^2 and χ^2 = Q) were also reported.

The results of the SMD test for quantitative features evaluated in at least two studies are reported in Table 6. Based on the SMD test, maximum dimension, ellipsoid volume, attenuation of pre-contrast liver, and attenuation of the liver in the arterial, portal, and delayed phase were considered informative. The SMD was negative for all informative quantitative CT features based on attenuation, meaning that malignant lesions showed lower attenuation values than benign lesions. Ellipsoid volume and the maximum dimension had positive SMD. Forest plots reporting the pooled SMD for the informative quantitative CT features are reported in Fig.3.

Table 6 Summary of meta-analysis for quantitative features

CT feature	**Studies (n)**	**Standardized mean difference**	**SE**	**t**	**p-value**
Maximum dimension	4	0.858	0.35	2.44	0.015
Ellipsoid volume	2	0.567	0.28	2.97	0.003
Attenuation of pre-contrast normal liver	4	-0.596	0.3	-2	0.046
Attenuation of pre-contrast lesion	3	-0.988	0.52	-1.88	0.062
Attenuation of arterial phase normal liver	3	-0.209	0.16	-1.34	0.183
Attenuation of arterial phase lesion	3	-0.517	0.16	-3.26	0.001
Attenuation of portal phase normal liver	3	-0.057	0.35	-0.16	0.869
Attenuation of portal phase lesion	3	-0.526	0.16	-3.31	0.001
Attenuation of delayed phase normal liver	3	-0.107	0.16	-0.68	0.497
Attenuation of delayed phase lesion	5	-0.377	0.13	-2.98	0.003

SE, standard error; t, Student’s t-test



Maximum Dimension (cm)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Burti et al., 2021	32	6.39	4.72	37	4.48	3.87	0.44	[-0.0; 0.9]	26.8%
Leela-Arporn et al., 2019	52	6.60	3.10	18	3.10	1.40	1.25	[0.7; 1.8]	25.4%
Jones et al., 2016	14	6.10	17.00	10	6.00	7.82	0.01	[-0.8; 0.8]	21.6%
Taniura et al., 2009	36	6.78	3.53	40	2.74	0.59	1.62	[1.1; 2.1]	26.2%
Random effects model	**134**			**105**			**0.86**	**[0.2; 1.6]**	**100.0%**

Heterogeneity: $I^2 = 82\%$, $\tau^2 = 0.40$, $p < 0.01$

Ellipsoid Volume (cm^3)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Burti et al., 2021	32	222.90	380.91	37	111.21	343.89	0.31	[-0.2; 0.8]	53.3%
Leela-Arporn et al., 2019	52	195.00	226.60	18	21.30	24.50	0.87	[0.3; 1.4]	46.7%
Random effects model	**84**			**55**			**0.57**	**[0.0; 1.1]**	**100.0%**

Heterogeneity: $I^2 = 57\%$, $\chi_1^2 = 2.31$ ($p = 0.13$)

Attenuation of pre-contrast normal liver (HU)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Stehlik et al., 2021	14	63.10	9.02	17	60.90	8.69	0.24	[-0.5; 1.0]	23.4%
Burti et al., 2021	32	59.39	8.32	37	65.26	7.99	-0.71	[-1.2; -0.2]	28.9%
Jones et al., 2016	14	60.00	5.50	10	71.00	7.50	-1.66	[-2.6; -0.7]	18.1%
Taniura et al., 2009	36	58.10	11.20	40	62.70	8.10	-0.47	[-0.9; 0.0]	29.7%
Random effects model	**96**			**104**			**-0.59**	**[-1.2; 0.0]**	**100.0%**

Heterogeneity: $I^2 = 71\%$, $\tau^2 = 0.23$, $p = 0.02$

Attenuation of lesion in arterial phase (HU)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Stehlik et al., 2021	14	52.80	17.24	17	65.00	28.80	-0.49	[-1.2; 0.2]	19.2%
Leela-Arporn et al., 2019	52	112.00	49.60	18	134.70	72.30	-0.40	[-0.9; 0.1]	34.1%
Taniura et al., 2009	36	91.10	35.20	40	116.10	44.40	-0.61	[-1.1; -0.2]	46.7%
Random effects model	**102**			**75**			**-0.52**	**[-0.8; -0.2]**	**100.0%**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.84$

Attenuation of lesion in portal phase (HU)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Stehlik et al., 2021	14	87.00	30.20	17	83.20	29.28	0.12	[-0.6; 0.8]	12.6%
Burti et al., 2021	32	89.62	40.60	37	90.13	41.96	-0.01	[-0.5; 0.5]	28.2%
Leela-Arporn et al., 2019	52	105.20	28.10	18	119.70	37.50	-0.47	[-1.0; 0.1]	21.5%
Jones et al., 2016	14	77.00	19.50	10	99.00	24.00	-0.99	[-1.9; -0.1]	8.4%
Taniura et al., 2009	36	104.90	34.60	40	123.90	17.50	-0.70	[-1.2; -0.2]	29.3%
Random effects model	**148**			**122**			**-0.38**	**[-0.6; -0.1]**	**100.0%**

Heterogeneity: $I^2 = 50\%$, $\tau^2 = 0.09$, $p = 0.09$

Attenuation of lesion in delayed phase (HU)

Study	Malignant			Benign			SMD	95%CI	Weight
	Total	Mean	SD	Total	Mean	SD			
Stehlik et al., 2021	14	87.00	30.20	17	83.20	29.28	0.12	[-0.6; 0.8]	12.6%
Burti et al., 2021	32	89.62	40.60	37	90.13	41.96	-0.01	[-0.5; 0.5]	28.2%
Leela-Arporn et al., 2019	52	105.20	28.10	18	119.70	37.50	-0.47	[-1.0; 0.1]	21.5%
Jones et al., 2016	14	77.00	19.50	10	99.00	24.00	-0.99	[-1.9; -0.1]	8.4%
Taniura et al., 2009	36	104.90	34.60	40	123.90	17.50	-0.70	[-1.2; -0.2]	29.3%
				122			**-0.38**	**[-0.6; -0.1]**	**100.0%**

Figure 3 Pooled standardised mean difference (SMD) for maximum dimensions, ellipsoid volume, attenuation of pre-contrast normal liver, and attenuation of lesion in arterial, portal, and delayed phase post-contrast. 95% CI, 95% confidence interval; SD, standard deviation; HU, Hounsfield unit.

DISCUSSION

The results of this meta-analysis revealed that, based on the available literature, some qualitative CT features showed statistically significant differences between benign and malignant FLL. The presence of a capsule, hypoattenuation, and heterogeneity in the delayed phase showed the highest DOR and, therefore, are the most reliable qualitative CT features for the detection of malignant FLL. Well-defined margins, abnormal lymph nodes, heterogeneity in the arterial phase, and heterogeneity in the portal phase were also found to be informative, albeit with a lower level of confidence. The quantitative features of maximal FLL dimension, attenuation of pre-contrast normal liver, and attenuation of the normal liver in the arterial, portal, and delayed phases had statistically significant differences between benign and malignant groups. Interestingly, the attenuation of the CT-normal liver parenchyma showed statistically significant differences between benign and malignant FLL. The finding of significant differences between attenuation of what was considered to be ‘normal’ liver between the two groups might indicate that the ‘normal’ liver parenchyma surrounding the malignant FLL could also be involved in the neoplastic process. This implies that, in addition to sampling the FLL, sampling of the ‘normal’ liver parenchyma should be considered.

As only individual CT features could be analysed in this meta-analysis, the overall accuracy of CT in the detection of malignant FLL could not be determined. However, the results provided in Tables 5 and 6 are, in the authors’ opinion, a valuable aid for the veterinary radiologist in characterising FLL in dogs, as they indicate several qualitative and quantitative CT features that might be useful in differentiating between malignant and benign lesions. However, the diagnostic accuracy of even the most discriminating CT features was only moderate such that aspirate cytology and, in some cases, biopsy histopathology remain necessary to accurately characterise FLL. Nonetheless, informative CT features can lend weight to the results of pathology and provide a more accurate evaluation of FLL in dogs, while some FLL are located at sites inaccessible to sampling.

Another aspect that emerged from this meta-analysis was the relatively low number of studies (and cases) available in the veterinary literature on this topic, especially when compared to the human literature. The CT features of only 419 FLL in dogs had been described, whereas meta-analyses including thousands of patients are currently available in the medical literature (Lee et al., 2015; Roberts et al., 2018a). Such a dissimilarity is, most likely, related to differences in the standards of care between humans and dogs both in access to CT imaging and to subsequent investigation of FLL identified. It is anticipated that, as the number of CT scanners available in veterinary clinics increases, a higher number of cases and FLL will be available for review.

A limitation of this meta-analysis was related to the different scanning protocols used in the included studies. Burti et al. (2021) and Jones et al. (2016) used a dual-phase CT scanning protocol that included only a delayed phase post-contrast scan, while the remaining studies used a triple-phase CT scanning protocol that included arterial, portal, and delayed phase post-contrast scans. Most of the CT features that were deemed informative from the meta-analysis results can be evaluated using either CT scanning protocol. Both heterogeneity and attenuation were informative in all phases, and the remaining informative CT features (i.e. well-defined margins, presence of a capsule, lymph nodes, and maximum dimension) could be evaluated independent of the scanning protocol. Due to the limited number of included studies, a subgroup analysis was not performed.

A low risk of bias was attributed to most of the considered studies. While this indicated an acceptable quality of the veterinary literature on this topic, it did highlight some frequently encountered limitations. For example, none of the authors specified whether the pathologists performing the cytological or histopathological analyses were blinded to the results of the CT imaging. It is suspected that the pathologists had not be blinded, as many of the studies were retrospective and during routine clinical practice there is constant dialogue between radiologists and pathologists, as this is considered to enhance diagnostic accuracy (Raab et al., 2000). The retrospective nature of many of the studies may also have of rendered subsequent blinded review of the cytological or histopathological samples unfeasible. The risk of bias for case selection was considered as high in two studies, Taniura et al. (2009) and Kutara et al. (2014), as they limited their inclusion criteria to only certain FLL histotypes. Neither paper provided justification for such a choice, and therefore a selection bias was evident. The inclusion of data from these studies might have influenced the overall results of the meta-analysis towards an increased pooled diagnostic accuracy for some of the CT features. It is warranted that future studies on this topic should adopt study designs that avoid selection bias.

None of the studies included a power analysis or equivalent estimation method. A lack of effectiveness in study design, as previously highlighted by Di Girolamo and Reynders (2016), is a general problem in the veterinary medical literature. Low incidence rates of primary hepatic neoplasms in dogs (Marolf, 2017), along with the increasing publication pressure exerted on researchers, make such a topic attractive for retrospective, rather than prospective, study.

Because HCC is the most common histotype of FLL in humans and is the liver neoplasia with the most available treatment options (Lee et al., 2015), the medical literature on FLL is mainly focused on the detection and grading of HCC. A continuously updating, extensive and detailed algorithm for the imaging, reporting, and care of HCC in humans, the Liver Imaging Reporting and Data System (LI-RADS; Elsayes et al., 2017) contains guidelines for the use of various diagnostic imaging modalities, including CT, in the surveillance and grading of HCC. In humans, HCC mainly arise in cirrhotic livers, while the correlation between hepatic degenerative disease and HCC has not yet been demonstrated in dogs. The efficacy of some treatments for HCCs have been studied in dogs (Marconato et al., 2020), while treatments for the other liver malignancies have been scarcely described. A veterinary counterpart of the human LI-RADS is currently not available, in part due to a lack of meta-analyses on this topic. Due to the relatively low number of studies available in the literature, the specific CT features of HCC in dogs could not be evaluated separately to those of other malignant FLL in the present meta-analysis. However, increasing numbers of publications focusing on the description of the diagnostic imaging features of FLL in dogs have become available in the last few years. In particular, the ultrasonographic (Warren-Smith et al., 2012), contrast-enhanced ultrasonographic (Nakamura et al., 2010; Morishita et al., 2017; Banzato et al., 2019; Burti et al., 2020), CT, and MRI (Constant et al., 2016; Borusewicz et al., 2019) features of FLL (including HCC) have been described. The data reported in the present study point towards the possible creation of diagnostic algorithms for the diagnosis and management of FLL in dogs.

Most of the meta-analyses on FLL in the human medical literature focus on the relative diagnostic accuracy of MRI and CT in the detection of HCC (Lee et al., 2015; Roberts et al., 2018b). A comparison of the diagnostic accuracy of MRI and CT in differentiating between benign and malignant FLL in dogs is not feasible at the present time, as the MRI features of FLL in dogs have rarely been evaluated (Constant et al., 2016; Borusewicz et al., 2019). The low number of publications on this topic is, most likely, related to the limited availability of high-field MRI scanners in the veterinary sector and their limited use in evaluating the abdomen.

A final limitation of the present meta-analysis is that techniques to account for the non-independence of sensitivity and specificity, such as bivariate models and hierarchical summary ROCS could not be performed due to the low number of studies that fit the inclusion criteria (Harbor et al., 2007). Additional studies on this topic are required to remedy this.

CONCLUSIONS

The qualitative and quantitative CT features to differentiate malignant from benign FLL were analysed. Well-defined margins, presence of a capsule, abnormal lymph nodes, heterogeneity in the arterial, portal, and delayed phase, maximum dimension, ellipsoid volume, attenuation of the liver in the pre-contrast phase, and attenuation of the liver in the arterial, portal, and delayed phase were found to be informative.

Conflict of interest statement None of the authors has any other financial or personal relationships that could inappropriately influence or bias the content of the paper.

Acknowledgements The present paper is part of a project funded by a research grant from the Department of Animal Medicine, Production and Health – MAPS, University of Padua, Italy: SID- Zotti 2018 (€ 32,000; Application of deep-learning algorithms in pet animal diagnostic imaging)

Supplementary material Supplementary data associated with this article can be found, in the online version, at doi: https://doi.org/10.1016/j.tvjl.2021.105773

BIBLIOGRAPHY

1. Ariff, B., Lloyd, C.R., Khan, S., Shariff, M., Thillainayagan, A.V., Bansi, D.S., Khan, S.A., Taylor-Robinson, S.D., Lim, A.K.P., 2009. Imaging of liver cancer. World Journal of Gastroenterology 15, 1289-1300.
2. Banzato, T., Rubini, G., Orlandi, R., Bargellini, P., Bonsembiante, F., Zotti, A., 2019. Contrast enhanced ultrasound fetaures of hepatocellular carcinoma in dogs. Veterinary Record 186, 187.
3. Borusewicz, P., Stańczyk, E., Kubiak, K., Spuzak, J., Glińska-Suchocka, K., Jankowski, M., Sławuta, P., Kubiak-Nowak, D., Podgórski, P., 2019. Magnetic resonance imaging of liver tumors using gadoxetic acid (Gd-EOB-DTPA) - Pilot study. BMC Veterinary Research 15, 1-9.
4. Burti, S., Zotti, A., Bonsembiante, F., Contiero, B., Banzato, T., 2021. Diagnostic accuracy of delayed phase post contrast computed tomographic images in the diagnosis of focal liver lesions in dogs: 69 Cases. Frontiers in Veterinary Science 8, 1-10.
5. Burti, S., Zotti, A., Rubini, G., Orlandi, R., Bargellini, P., Bonsembiante, F., Banzato, T., 2020. Contrast-enhanced ultrasound features of malignant focal liver masses in dogs. Scientific Reports 10, 1-12.
6. Chou, R., Cuevas, C., Fu, R., Devine, B., Wasson, N., Ginsburg, A., Zakher, B., Pappas, M., Graham, E., Sullivan, S.D, 2015. Imaging techniques for the diagnosis of hepatocellular
7. carcinoma: A systematic review and meta-analysis. Annals of Internal Medicine 162, 697-711.
8. Constant, C., Hecht, S., Craig, L.E., Lux, C.N., Cannon, C.M., Conklin, G.A., 2016. Gadoxetate disodium (gd-eob-dtpa) contrast enhanced magnetic resonance imaging characteristics of hepatocellular carcinoma in dogs. Veterinary Radiology and Ultrasound 57, 594-600.
9. DerSimonian, R., Laird, N., 2015. Meta-analysis in clinical trials revisited. Contemporary Clinical Trials 45, 139-145.
10. Di Girolamo, N., Reynders, R.M., 2016. Deficiencies of effectiveness of intervention studies in veterinary medicine: A cross-sectional survey of ten leading veterinary and medical journals. Peer J 2016, 1-22.
11. Elsayes, K.M., Kielar, A.Z., Agrons, M.M., Szklaruk, J., Tang, A., Bashir, M.R., Mitchell, D.G., Do, R.K., Fowler, K.J., Chernyak, V., et al., 2017. liver imaging reporting and data system: An expert consensus statement. Journal of Hepatocellular Carcinoma Volume 4, 29-39.
12. Fukushima, K., Kanemoto, H., Ohno, K., Takahashi, M., Nakashima, K., Fujino, Y., Uchida, K., Fujiwara, R., Nishimura, R., Tsujimoto, H., 2012. CT characteristics of primary hepatic mass lesions in dogs. Veterinary Radiology and Ultrasound 53, 252-257.

13. Griebie, E.R., David, F.H., Ober, C.P., Feeney, D.A., Anderson, K.L., Wuenschmann, A., Jessen, C.R., 2017. Evaluation of canine hepatic masses by use of triphasic computed tomography and B-mode, color flow, power, and pulsed-wave Doppler ultrasonography and correlation with histopathologic classification. American Journal of Veterinary Research 78, 1273-1283.
14. Higgins, J.P.T., 2003. Measuring inconsistency in meta-analyses. British Medical Journal 327, 557-560.
15. Harbord, R.M., Deeks, J.J., Egger, M., Whiting, P., Sterne, J.C., 2007 A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 8, 239-251.
16. Jones, I.D., Lamb, C.R., Drees, R., Priestnall, S.L., Mantis, P., 2016. Associations between dual-phase computed tomography features and histopathologic diagnoses in 52 dogs with hepatic or splenic masses. Veterinary Radiolology and Ultrasound 57, 144-153.
17. Kutara, K., Asano, K., Kito, A., Teshima, K., Kato, Y., Sasaki, Y., Edamura, K., Shibuya, H., Sato, T., Hasegawa, A., et al., 2006. Contrast harmonic imaging of canine hepatic tumors. Journal of Veterinary Medical Science 68, 433-438.
18. Kutara, K., Seki, M., Ishikawa, C., Sakai, M., Kagawa, Y., Iida, G., Ishigaki, K., Teshima, K., Edamura, K., Nakayama, T., et al., 2014. Triple-phase helical computed tomography in dogs with hepatic masses. Veterinary Radiology and Ultrasound 55, 7-15.
19. Lee, Y.J., Lee, J.M., Lee, J.S., Lee, H.Y., Park, B.H., Kim, Y.H., Han, J.K., Choi, B.I., 2015. Hepatocellular carcinoma: Diagnostic performance of multidetector CT and MR imaging-a systematic review and meta-analysis. Radiology 275, 97-109.
20. Lee, S., Kim, S.S., Roh, Y.H., Choi, J.Y., Park, M.S., Kim, M.J., 2020. Diagnostic performance of CT/MRI liver imaging reporting and data system 2017 for hepatocellular carcinoma: A systematic review and meta-analysis. Liver International 40, 1488-1497.
21. Marconato, L., Sabattini, S., Marisi, G., Rossi, F., Leone, V.F., Casadei-Gardini, A., 2020. Sorafenib for the treatment of unresectable hepatocellular carcinoma: Preliminary toxicity and activity data in dogs. Cancers 12, 1272.
22. Marolf, A.J., 2017. Diagnostic imaging of the hepatobiliary system: An update. Veterinary Clinics of North America - Small Animal Pract. 47, 555-568.
23. Methley, A.M., Campbell, S., Chew-Graham, C., McNally, R., Cheraghi-Sohi, S., 2014. PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Service Reserch 14.
24. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Altman, D., Antes, G., Atkins, D., Barbour, V.,

Barrowman, N., Berlin, J.A., et al., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine 6.

25. Morishita, K., Hiramoto, A., Michishita, A., Takagi, S., Osuga, T., Lim, S.Y., Nakamura, K., Sasaki, N., Ohta, H., Takiguchi, M., 2017. Washout ratio in the hepatic vein measured by contrast-enhanced ultrasonography to distinguish between inflammatory and noninflammatory hepatic disorders in dogs. Journal of Veterinary Internal Medicine 31, 770-777.
26. Nakamura, K., Takagi, S., Sasaki, N., Bandula Kumara, W.R., Murakami, M., Ohta, H., Yamasaki, M., Takiguchi, M., 2010. Contrast-enhanced ultrasonography for characterization of canine focal liver lesions. Veterinary Radiology and Ultrasound 51, 79-85.
27. Raab, S.S., Oweity, T., Hughes, J.H., Salomao, D.R., Kelley, C.M., Flynn, C.M., D'Antonio, J.A., Cohen, M.B., 2000 Effect of clinical history on diagnostic accuracy in the cytologic interpretation of bronchial brush specimens. American Journal of Clinical Pathology 114,78-83.
28. Roberts, L.R., Sirlin, C.B., Zaiem, F., Almasri, J., Prokop, L.J., Heimbach, J.K., Murad, M.H., Mohammed, K., 2018. Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis. Hepatology 67, 401-421.
29. Shah, S., Shukla, A., Paunipagar, B., 2014. Radiological features of hepatocellular carcinoma. Journal of Clinical and Experimental Hepatology 4, S63-S66.
30. Shim, S.R., Kim, S.J., Lee, J., Rücker, G., 2019 Network meta-analysis: Application and practice using R software. Epidemiological Health 4.
31. Stehlík, L., Di Tommaso, M., Del Signore, F., Paninárová, M., Terragni, R., Magni, T., Pontonutti, L., Carloni, A., Alberti, M., De Magistris, A. V., et al., 2020. Triple-phase multidetector computed tomography in distinguishing canine hepatic lesions. Animals 11, 11.
32. Taniura, T., Marukawa, K., Yamada, K., Hikasa, Y., Ito, K., 2009. Differential diagnosis of hepatic tumor-like lesions in dog by using dynamic CT scanning. Hiroshima Journal of Medical Science 58, 17-24.
33. Warren-Smith, C.M.R., Andrew, S., Mantis, P., Lamb, C.R., 2012. Lack of associations between ultrasonographic appearance of parenchymal lesions of the canine liver and histological diagnosis. Journal of Small Animal Practice 53, 168-173.
34. Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B.,
35. Leeflang, M.M., Sterne, J.A., Bossuyt, P.M., 2011. QUADAS-2 Group. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine 155, 529-36.

CHAPTER FOUR

A machine learning-based approach for classification of focal splenic lesions based on their CT features.

The chapter was adapted from:

Burti S, Zotti A, Bonsembiante F, Contiero B and Banzato T (2022) *A Machine Learning-Based Approach for Classification of Focal Splenic Lesions Based on Their CT Features.* Front. Vet. Sci. 9:872618. doi: 10.3389/fvets.2022.872618

ABSTRACT

The aim of the study was to describe the CT features of focal splenic lesions (FSLs) in dogs in order to predict lesion histotype. Dogs that underwent a CT scan and had a FSL diagnosis by cytology or histopathology were retrospectively included in the study. For the statistical analysis the cases were divided into four groups, based on the results of cytopatholoy or hystopathology, namely: nodular hyperplasia (NH), other benign lesions (OBLs), sarcoma (SA), round cell tumour (RCT). Several qualitative and quantitative CT features were described for each case. The relationship occurring between each individual CT feature and the histopathological groups was explored by means of c chi-square test for the count data and by means of Kruskal-Wallis or ANOVA for the continuous data. Furthermore, the main features of each group were described using factorial discriminant analysis, and a decision tree for lesion classification was then developed. Sarcomas were characterised by large dimensions, a cystic appearance and an overall low post contrast-enhancement. NH and OBLs were characterised by small dimensions, a solid appearance and a high post-contrast enhancement. OBLs showed higher post-contrast values than NH. Lastly, RCTs did not exhibit any distinctive CT features. The proposed decision tree had a high accuracy for the classification of SA (0.89) and a moderate accuracy for the classification of OBLs and NH (0.79), whereas it was unable to classify RCTs. The results of the factorial analysis and the proposed decision tree could help the clinician in classifying FSLs based on their CT features. A definitive FSL diagnosis can only be obtained by microscopic examination of the spleen.

INTRODUCTION

Focal splenic lesions (FSLs) are common in dogs, especially in elderly subjects. Most FSLs (51%) are benign; the most common histotypes are haematoma, nodular hyperplasia, and myelolipoma (1–3). Haemangiosarcoma is reported as the most common primary malignant tumour of the spleen, accounting for almost 80% of malignant FSLs (1,2), followed by fibrosarcoma and leiomyosarcoma. Splenic metastases (from other primary sarcomas, carcinomas or neuroendocrine tumours in most cases) are less common, accounting for 1-6% of the total of the FSLs (3).

Despite FSLs being a common finding in canine ultrasound (US) and computed tomography (CT) (4) there is a general paucity of studies systematically describing their imaging features. No specific US features are reported as useful in distinguishing between different FSL histotypes (5). Previous studies describing the CT features of FSLs have reported conflicting results. Fife et al., 2004 (6), reported that, in dual-phase CT imaging, a FSL with a post-contrast Hounsfield Unit value lower than 55 is most likely malignant. However, Jones et al., 2016 (7), reported no dual-phase CT features as useful in the distinction between benign and malignant lesions. Kutara et al., 2017(2), using triple-phase CT imaging, reported some CT features (lesion volume and homogeneous contrast enhancement) as useful in differentiating between haematoma, nodular hyperplasia, haemangiosarcoma and undifferentiated sarcoma in dogs. Lastly, Lee et al, 2018 (8) reported triple-phase CT, combined with ultrasonography, as useful in the differentiation between benign and malignant lesions.

In the last few years, an increasing number of research papers exploring the possible applications of machine learning in veterinary radiology have been published (9-15). Research in this field has mostly been focused on the automatic classification of radiographic images (14,16,17), the distinction between benign and malignant brain lesions on MRI (10,18), and the classification of liver focal lesion types on CT images (19). To the best of the authors’ knowledge, the approach of applying machine learning to classify splenic lesions based on their CT appearance has not yet been explored.

In such a scenario, the aims of this study are: 1) to describe the CT features of FSLs in dogs; 2) to use machine learning algorithms to describe the complex relationship existing between different FSL histotypes and their CT features; and 3) to develop an easy-to-use algorithm for classifying FSLs based on their CT features.

MATERIAL AND METHODS

Study Population

The medical records of 62 dogs (32 males and 30 females – mean age 10.4 ± 2.3 years) referred to the Pedrani Veterinary Clinic (Via Caldierino 14, Zugliano, Vicenza, Italy) and to the Veterinary Teaching Hospital of the University of Padua (Viale Dell'Università 16, Legnaro, Padua, Italy) between June 2015 and November 2021 were prospectively collected. Criteria for inclusion in the study were: 1) a CT scan was conducted, 2) cytopathological and/or histopathological diagnosis of the splenic lesion. Exclusion criteria were: 1) chemotherapy at the time of the CT scan; 2) non-diagnostic cytopathological samples or equivocal cytopathological diagnosis. Patient signalment was recorded for each animal. The dogs belonged to several different breeds (31 mixed breeds, 4 Labrador Retrievers, 3 Golden Retrievers, 2 Boxers, 2 Bernese Mountain dogs, 2 German Shepherds, 2 Cockers, 2 Cane Corso, and one each of Fox Terrier, Yorkshire Terrier, English setter, Whippet, Great Dane, Weimaraner, Pointer, Jack Russell Terrier, Belgian Shepherd Dog, Australian Shepherd Dog, Shih Tzu, Lakeland Terrier and Hovawart). Six dogs were excluded because they were receiving chemotherapy at the time of the CT scan, and four were excluded because the cytopathological samples resulted as non-diagnostic. Of the remaining 52 dogs, 16 had a final diagnosis of nodular hyperplasia, 6 of normal splenic parenchyma, 5 of extramedullary haematopoiesis, 3 of haematoma, 2 of lymphoma, 2 of histiocytic sarcoma, 2 of mastocytoma, 1 of mesenchymal neoplasia, 1 of plasma-cell neoplasia, and 14 of sarcoma (5 sarcoma, 4 stromal sarcoma, 3 hemangiosarcoma, 1 leiomyosarcoma, and 1 myxoid liposarcoma). The cases were grouped into four broader histological categories for the statistical analysis: nodular hyperplasia NH, 16 cases; other benign lesions (OBLs), 14 cases; round cell tumour (RCT), 8 cases, sarcoma (SA), 14 cases.

All the methods were carried out in compliance with the relevant guidelines and regulations. This study was conducted respecting the Italian Legislative Decree N° 26/2014 (transposing EU Directive 2010/63/EU). Nevertheless, since the data used in this study were part of routine clinical activity, no ethical committee approval was required. Informed consent for personal data processing was obtained from the owners.

Cytopathological and Histopathological Examination

Thirty-three splenic masses were sampled through ultrasound-guided fine needle aspiration for cytological assessment. 21-gauge needles were always used. Cytological slides were obtained by smearing the aspirates on glass slides, which were subsequently air-dried, stained with May-Grünwald-Giemsa stain and cover-slipped. All the cytological slides were evaluated by the same cytologist (FB). Cytology was always performed immediately after the CT scan. Histopathology was not performed in any of these cases.

Twenty-one splenic masses were sampled through ultrasound-guided Tru-cut biopsy for histological assessment. Formalin-fixed tissue samples were dehydrated in a graded ethanol series and embedded in paraffin. 4-µm-thick sections were stained with haematoxylin and eosin and evaluated by one pathologist.

Computed Tomography Examination

Three different scanners were used to perform the CT examinations: Asteion super 4 (Toshiba Medical System Corporation), at the Veterinary Teaching Hospital; Revolution ACT, General Electric Medical System), and Optima CT 520 Series (General Electric Medical System) at the Pedrani Veterinary Clinic. The scanning protocols were slightly different for the different scanners. The protocols for the Asteion super 4 were: helical acquisition mode, exposure time of 0.725 seconds, voltage of 120 kV, amperage of 150 mA, and slice thickness of 1-3 mm. For the Revolution ACT, were: exposure time of 0.725 seconds, voltage of 100 kV, amperage of 100 mA, and slice thickness of 1-2.5 mm. Lastly for the Optima CT 520 Series were: exposure time of 0.725 seconds, voltage of 120 kV, amperage of 180 mA, and slice thickness of 1-3 mm.

All the dogs underwent a 12-hour fasting period prior to examination. All the examinations were performed on anaesthetised subjects placed in ventral recumbency. Contrast medium (Ioversol 350mg/ml, Optiray 350, Liebel-Flarsheim Company LLC, USA) was administered at the dosage of 660 mg/kg through two different modalities depending on the facility: 1) via an injector at the Pedrani Veterinary Clinic; 2) manually injected intravenously as a bolus at the Veterinary Teaching Hospital. Pre-contrast and delayed phase CT scans, the latter starting 30-50 seconds after the end of the contrast medium injection, were always performed.

All the images were stored as digital imaging and communication in medicine (DICOM) files.

Image Analysis

All the scans were retrieved using a picture archiving and communication system (PACS). All the images were reviewed in a soft tissue window (WW: 400 HU - WL: 40HU) using a commercially available software (Horos v3.3.6). In the case of multiple lesions in the same dog, the largest sampled lesion was described.

The following qualitative features were evaluated: 1) margins (well- or ill-defined); 2) surface (regular or irregular); 3) appearance (solid or cyst-like) - the lesion was classified as “cyst-like” in the presence of at least one area with a measured Hounsfield Unit (HU) value similar to that of the animal’s gallbladder (representing possible necrosis or haemorrhage) (20); 4) splenic lymph-nodes appearance (normal or abnormal) – splenic lymph nodes were classified as abnormal if any of the following changes were evident: a) lymphadenomegaly (the dimensions of the splenic lymph nodes were subjectively compared to the surrounding abdominal lymph nodes), b) heterogeneous appearance c) round or irregular shape; 5) homogeneity of contrast-medium distribution inside the lesion (homogeneous or heterogeneous); 6) enhancement pattern (prevalently central, rim enhancement, or diffuse distribution).

The following quantitative characteristics were evaluated: 1) attenuation (measured as an HU value) of the tomographically normal splenic parenchyma, in both the pre-contrast and the delayed phase; 2) attenuation (mean HU value) of the lesion in both the pre-contrast and the delayed phase; 3) maximum transverse diameter; 4) volume - the shape of the lesion was considered to be an ellipsoid and the formula $V = \frac{4}{3}\pi\,(\mathrm{height}/2 * \mathrm{width}/2 * \mathrm{length}/2)$ was applied (21); 5) attenuation of the lesion compared to that of the radiologically normal splenic parenchyma in the pre-contrast images (hypoattenuating, isoattenuating or hyperattenuating); 6) enhancement degree of the lesion compared to that of the radiologically normal splenic parenchyma in post-contrast images (hypoenhancing, isoenhancing or hyperenhancing). The attenuation and the enhancement degree of the lesion were determined based on the difference between the mean HU value measured on the lesion and the HU value measured on the radiologically normal splenic parenchyma. The lesions were classified as: a) isoattenuating/isoenhancing if the difference fell in the ±10 HU range; b) hyperattenuating/hyperenhancing with a difference greater than +10 HU; c) hypoattenuating/hypoenhancing if the difference was lower than -10 HU (7,22). The HU values were measured in three circular regions of interest (ROIs), in both the normal and in the pathological parenchyma, carefully avoiding cystic regions and vascular structures. The same ROIs were selected in pre- and post-contrast images. The size of the ROI was manually adjusted for each case.

The CT features were evaluated separately by two of the authors of this study (SB: with 4 years' experience in diagnostic imaging and AZ, with 20 years' experience in diagnostic imaging). The reviewers were blinded to the results of the histopathological examination.

Statistical Analysis

To compare the differences between the four diagnostic categories, the count data expressed as percentages were analysed with a chi-square test (or Fisher's exact test when there were fewer than 5 units of data). The quantitative variables were assessed for normality using Shapiro-Wilks test. Differences between the four diagnostic categories were analysed with a one-way analysis of variance (ANOVA) for normally distributed data, whereas the non-parametric Kruskal-Wallis test was used for non-normally distributed data. A Bonferroni post-hoc pairwise comparison test was performed. A $p<0.05$ was considered as statistically significant. The analyses were conducted with SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

To describe the complex relationship existing between all the different CT features and the histopathological groups, two different supervised machine learning techniques were applied. The first to be used was a dimensionality reduction technique, known as factorial discriminant analysis (FDA). This technique was chosen in order to identify which of the CT features best discriminated between the four histopathological categories. FDA aims to identify different linear combinations of original features (components - F) that provide the best possible separation of two or more classes of units. A coefficient is assigned to each original variable based on its relative ability to discriminate between different groups. Different components are computed and, usually, the first two components explain most of the variance in the dataset. The correlations between the original variables and components were calculated in our study and coefficient values of $|r|>0.6$ and >0.5 for the first component and the second component respectively were considered significant. Classification of all the cases based on the first two components is plotted on a Cartesian plane, where the position on the x-axis is determined by the results of F1 and the position on the y-axis is determined by the results of F2, and this enables the discrimination ability of the analysis to be visually assessed. Lastly, the centroids (i.e. the arithmetic mean positions of all the points in a group) are plotted. The further away each centroid is from the 0 of the Cartesian axes and from the centroids of the other groups, the better is the discrimination ability of the analysis is for that group. The factorial discriminant analysis was performed using XLStat (Addinsoft 2022, XLSTAT statistical and data analysis solution, New York, USA).

Decision tree analysis was then performed to detect the best discriminating CT features (a recursive partitioning method was adopted using the rpart package of R-https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf, and a three-step procedure was applied to build the decision tree: 1) the features that provided the best data splitting were selected; 2)10-fold cross-validation was used to prune the decision tree having the lowest number of branches and the lowest misclassification rate (23); 3) a confusion matrix was built by comparing the values of actual vs predicted samples (obtained from the decision tree classification), and some quality indices regarding model performance were calculated (sensitivity, specificity, accuracy and misclassification rate).

RESULTS

The results of the analysis of the qualitative and quantitative features of the images, along with their p-values, are reported in Table 1 and Table 2 respectively. Pre- and post- contrast example images for each histopathological category are reported in Figures 1-4. Among the qualitative features, only the surface (χ^2 =8.71; p-value=0.033) and the appearance (χ^2 =12.98; p-value=0.005) showed statistically significant differences between histopathological groups. In particular, the main differences were found between OBLs and SAs for both the surface and the appearance. In fact, almost all (13/14) the SAs had an irregular surface and a cyst-like appearance. Instead, OBLs showed mainly a solid appearance (11/14) whereas surface was almost evenly distributed between regular (8/14) and irregular (6/14). The margins (χ^2 =5.12; p-value=0.163), lymph nodes (p-value= 0.169), post contrast homogeneity (χ^2 =4.37; p-value=0.224), and enhancement pattern (χ^2 =1.10; p-value=0.776) did not show statistically significant differences between the histopathological groups.

Table 3 Qualitative features, along with cytological or histological classification.

	Category					
	Nodular hyperplasia*(n = 16)*	**Other benign lesions[†] *(n = 14)***	**Round cells tumors[+] *(n = 8)***	**Sarcoma[++] *(n = 14)***	**Total *(n = 52)***	**p-value**
Margins*						0.163
well defined	9 (56%)	10 (71%)	6 (75%)	13 (93%)	38 (73%)	
ill defined	7 (44%)	4 (29%)	2 (25%)	1 (7%)	14 (27%)	
Surface*						0.033
regular	7 (44%)	8 (57%)	2 (25%)	1 (7%)	18 (35%)	
irregular	9 (56%) [ab]	6 (43%) [b]	6 (75%) [ab]	13 (93%) [a]	34 (65%)	
Aspect*						0.005
solid	10 (62%)	11 (79%)	5 (62%)	2 (14%)	28 (54%)	
cyst-like	6 (38%) [ab]	3 (21%) [b]	3 (38%) [ab]	12 (86%) [a]	24 (46%)	
Lymph-nodes**						0.169
normal	13 (81%)	10 (71%)	5 (63%)	6 (43%)	34 (65%)	
abnormal	3 (19%)	4 (29%)	3 (38%)	8 (57%)	18 (35%)	
Homogeneity post-contrast medium*						0.224
homogeneous	5 (31%)	5 (36%)	1 (14%)	1 (7%)	12 (23%)	
heterogeneous	11 (69%)	9 (64%)	7 (86%)	13 (93%)	40 (77%)	
Enhacement pattern*						0.776
diffuse	12 (75%)	11 (79%)	5 (63%)	9 (64%)	37 (71%)	
marginal	4	3	3	5	15 (29%)	
central	0	0	0	0	0	
Attenuation pre-contrast medium**						0.171
hypoattenuating	7 (44%)	6 (43%)	5 (63%)	11 (79%)	29 (56%)	
isoattenuating	9 (56%)	6 (43%)	3 (38%)	3 (21%)	21 (40%)	
hyperattenuating	0	2 (14%)	0	0	2 (4%)	
Enhancement post-contrast medium**						0.309
hypoenhancing	8 (50%)	7 (50%)	5 (63%)	12 (86%)	32 (61%)	
isoenhancing	1 (6%)	1 (7%)	1 (13%)	0	3 (6%)	
hyperenhancing	7 (44%)	6 (43%)	2 (25%)	2 (14%)	17 (33%)	

Different letters along columns mean significant values for $p < 0.05$

** k proportion test*

*** Fisher's exact test*

† Other benign lesions = 6 normal parenchyma, 5 extramedullary hematopoiesis, 3 hematomas

+ Round cells tumors = 2 mastocytomas, 2 lymphomas, 2 hystiocitic sarcoma, 1 mesenchimal neoplasia, 1 plasmacellular neoplasia

++ Sarcoma = 5 sarcoma, 4 stromal sarcoma, 3 hemangiosarcoma, 1 leiomyosarcoma, 1 myxoid liposarcoma

Table 2 Quantitative features, along with the cytological or histological classification

	Category					
	Nodular hyperplasia ***(n = 16)***	**Other benign lesions** † ***(n = 14)***	**Round cells tumors** + ***(n = 8)***	**Sarcoma**++ ***(n = 14)***	**Total** ***(n = 52)***	**p-value**
Max dimension (cm)*	2.17 (1.65-2.97) ^{b}	5.09 (2.50-8.65) ^{ab}	5.47 (1.22-12.37) ^{ab}	10.67 (7.57-16.00) ^{a}	4.59 (1.96-10.28)	0.001
Ellipsoid volume (cm^3)*	2.77 (1.45-8.32) ^{b}	23.45 (5.50-315.99) ^{b}	72.20 (0.62-538.71) ^{b}	375.24 (152.03-1387.69)	23.45 (2.39-350.86)	0.001
HU normal spleen pre-contrast medium*	61.03±6.01	63.11±11.59	55.11±9.34	54.18±7.06	58.83±9.18	0.026
HU normal spleen post-contrast medium	106.29±17.05	108.61±24.31	116.65±20.88	108.46±14.74	109.09±19.05	0.665
HU lesion pre-contrast medium**	49.59±12.67 ^{ab}	60.75±24.59 ^{a}	42.49±10.91 ^{ab}	33.26±11.29 ^{b}	47.10±18.95	0.001
HU lesion post-contrast medium**	93.90±29.19 ^{ab}	106.77±47.24 ^{a}	88.85±31.28 ^{ab}	60.17±32.23 ^{b}	87.51±39.22	0.010

Different letters along columns mean significant values for p < 0.05
** Kruskal-Wallis test*
*** One-way ANOVA*
† Other benign lesions = 6 normal parenchyma, 5 extramedullary hematopoiesis, 3 hematomas
+ Round cells tumors = 2 mastocytomas, 2 lymphomas, 2 hystiocitic sarcoma, 1 mesenchimal neoplasia, 1 plasmacellular neoplasia
++ Sarcoma = 5 sarcoma, 4 stromal sarcoma, 3 hemangiosarcoma, 1 leiomyosarcoma, 1 myxoid liposarcoma
HU = Hounsfield Unit

Maximum dimension and ellipsoid volume showed a non-normal distribution and, therefore, differences between the groups were calculated with the Kruskal-Wallis test. All the remaining variables showed a normal distribution and, therefore, differences were evaluated with the ANOVA. Most of the quantitative features revealed significant differences between the groups: HU value of pre-contrast normal spleen (F=3.37; *p*-value = 0.026), HU value of pre-contrast lesion (F=6.97; *p*-value= 0.001), HU value of post-contrast lesion (F=4.20; *p*-value=0.01), maximum dimension (k=16.13; *p*-value=0.001), and ellipsoid volume (k=16.94; *p*-value=0.001). Only the HU value of the post-contrast normal spleen showed no statistically significant differences between groups (F=0.53; *p*-value=0.665). Box-plots of all the quantitative variables are reported in Figure 5. It seems clear from analysis of the box plots that differences are mainly evident between sarcomas and other lesions. In particular the only statistically significant differences in the HU values of pre-contrast lesions and of post-contrast lesions are between OBLs and SAs. Only differences between NH and SAs were evident for both maximum dimension and ellipsoid volume (two highly correlated values).

The first two main components of the FDA (called F1 and F2) together explained about 86% of the total variability. The coefficients for F1 and F2 are reported in Table 3. The first component, explaining 63.82% of the total variability, is positively correlated (|r|>0.6) mainly with the HU value of the pre- and post-contrast lesion and with a solid appearance of the lesion, and is inversely correlated with maximum dimension and cystic appearance. The second component, explaining only 22% of the total variability, is moderately related only to pre-contrast hyperattenuatuation (|r|>0.5). Case distribution using the Cartesian system, based on classification by the two main components, is represented in Figure 6. From the graph in Figure 6a, and from the positions of the centroids (Figure 6b), it appears evident that the sarcoma group lies (almost completely) in the negative part of the x-axis, and is therefore associated with characteristics such as larger maximum dimensions and a cyst like appearance (that had a negative correlation to F1). The NH and OBL groups are both positioned in the positive part of the x-axis, and are therefore mainly characterised by smaller dimensions, higher pre- and post-contrast mean HU values and a solid appearance (positive correlation with F1). Furthermore, OBLs and NH are separated on the y-axis (F2), with OBLs exhibiting higher values than nodular hyperplasia. Lastly, RCTs are located in the centre of the Cartesian axis system and thus do not show any distinctive CT feature. Lastly, although an overall tendency for each group is noted, the large superimposition of the cases around the 0 on the Cartesian axes indicates that the subdivision of SA, OBL, and NH based on the CT features is suboptimal.

Table 3 F1 and F2 values of the factorial discriminant analysis based on 8 qualitative and 5 quantitative predictors

	F1	**F2**
Well-defined margins	-0.384	0.255
Ill-defined margins	0.384	-0.255
Regular surfaces	0.54	0.189
Irregular surfaces	-0.54	-0.189
Solid aspect	0.636	0.144
Cyst-like aspect	-0.636	-0.144
Normal lymph-nodes	0.413	-0.161
Abnormal lymph-nodes	-0.413	0.161
Homogeneous distribution	0.381	0.089
Heterogeneous distribution	-0.381	-0.089
Diffused contrast enhancement pattern	0.174	0.075
Marginal contrast enhancement pattern	-0.174	-0.075
Hypoattenuation	-0.421	-0.006
Isoattenuation	0.338	-0.202
Hyperattenuation	0.225	0.528
Hypoenhancement	-0.426	0.028
Isoenhancement	0.143	-0.035
Hyperenhancement	0.371	-0.011
Max dimension	-0.769	0.38
HU normal spleen pre-contrast medium	0.533	0.191
HU normal spleen post-contrast medium	-0.064	0.018
HU lesion pre-contrast medium	0.69	0.401
HU lesion post-contrast medium	0.593	0.178

The decision tree resulting from the analysis is reported in Figure 7. Three variables (max dimension, mean HU value of the pre-contrast lesion and HU value of the post-contrast lesion) were used for classification. Following the first split (max. dimensions < 3.6 cm), the decision tree classified 48% of the cases as nodular hyperplasia (max. dimension < 3.6 cm) and 52% of the cases as sarcoma (max dimension > 3.6 cm). Of these 48% classified as NH 56% were actually NH, 24% were OBLs, 16% were RCTs and only 4% were SAs. Instead, of the 52% of the cases classified as SA 7% were NH, 30% were OBLs, 15% were RCTs, and 48% were actually SAs. Following the second split on the left (mean HU value of the post-contrast lesion < 126), 35% of the cases were classified as NH (67% actually NH, 11% actually OBL, 17% actually RCT, 6% actually SA) and 13% (HU lesion post mean > 126) were classified as OBL (29% actually NH, 57% actually OBL, 14% actually RCT, and 0% actually SA). On another decision tree branch (mean HU values of pre-contrast lesion ≥44), 23% of the cases were classified as OBLs (17% actually NH, 58% actually OBL, 17% actually RCT, and 8% actually SA). On another secondary branch the algorithm classified 29% of the cases as sarcomas (0% actually NH, 7% actually OBL, 13% RCT, 80% actually SA).

Therefore, the following observations summarise the main findings of the decision tree. If the lesion is smaller than 3.6 cm and has a post-contrast HU value lower than 126 there, is a 67% chance that is NH (a 78% combined chance that it is benign if classing NH and OBLs together). If the lesion is smaller than 3.6 cm and has a post-contrast HU value higher than 126, there is a 57% chance that it is an OBL (an 86% combined chance that it is benign if classing OBLs and NH together). Instead, if the lesion is larger than 3.6 cm and has a mean pre-contrast mean HU value higher or equal to 44, there is a 58% chance the lesion is an OBL and a cumulative 75% chance it is benign. Lastly, if the lesion is larger than 3.6 cm and has a pre-contrast HU lower than 44, there is an 80% chance it is a SA and a cumulative 93% chance it is malignant (classing RCT and SA together). Not surprisingly, the algorithm did not identify any specific feature enabling the differentiation of RCTs from the remaining histopathological categories. The overall accuracy of the decision tree, when reapplied on the original data, was 0.67 and the *k* was 0.54. The sensitivity, the specificity and the balanced accuracy of the decision tree for each FSL category is reported in Tables 4-5.

Table 4 Confusion matrix that summarize the performance of the machine-learning based decision tree, giving the number of predicted cases

		Actual				
		Nodular hyperplasia	Other benign lesions	Round cells tumors	Sarcoma	**Total**
Predicted	Nodular hyperplasia	12	2	3	1	18
	Other benign lesions	4	11	3	1	19
	Round cells tumors	0	0	0	0	0
	Sarcoma	0	1	2	12	15
	Total	16	14	8	14	52

Table 5 Results of the classification of the FSLs based on the machine-learning based decision tree

	Nodular hyperplasia *(n = 16)*	**Other benign lesions *(n = 14)***	**Round cells tumors *(n = 8)***	**Sarcoma *(n = 14)***
Sensitivity	0.75	0.79	0.00	0.86
Specificity	0.83	0.79	1.00	0.92
Balanced accuracy	0.79	0.79	0.50	0.89

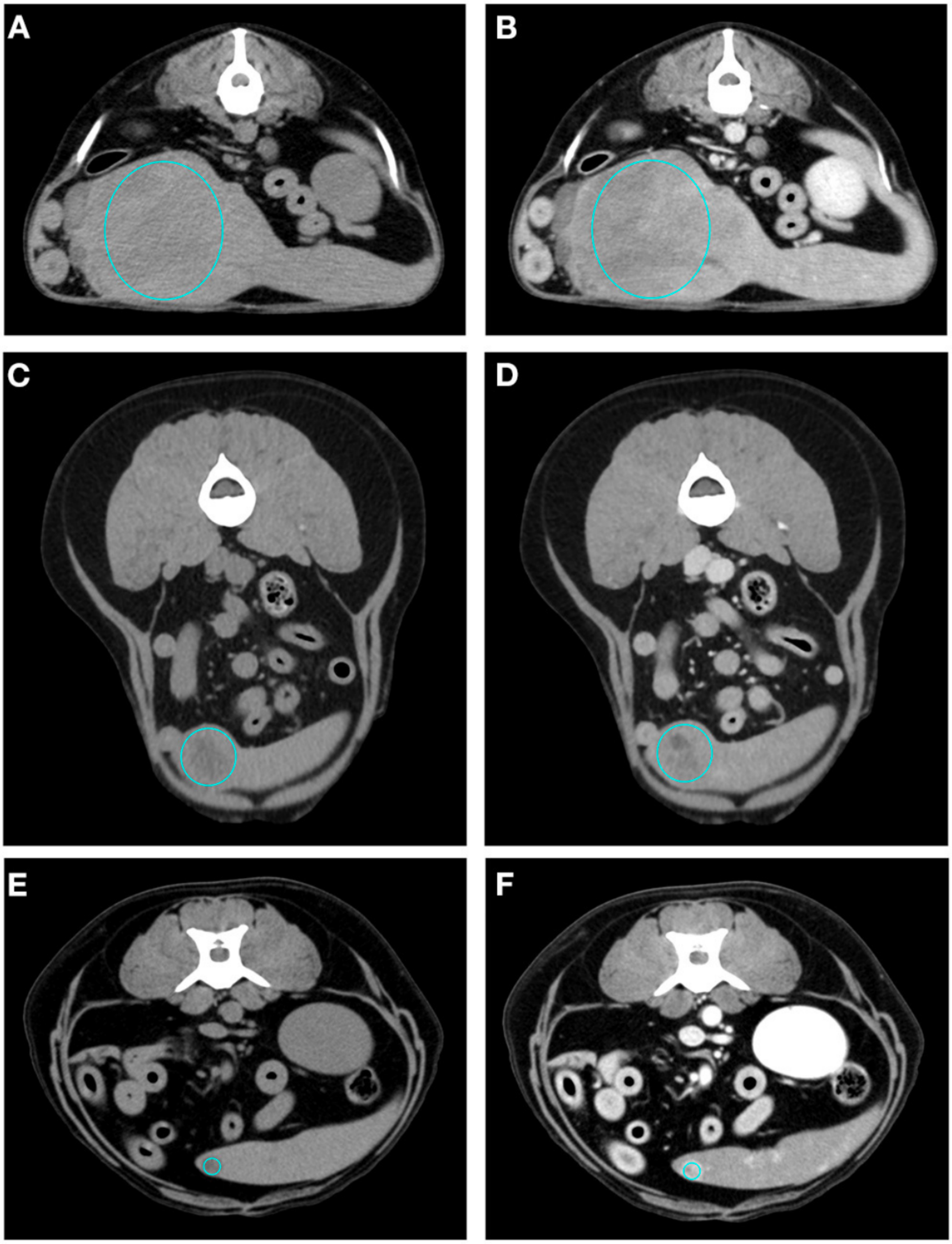


Figure 1 Pre- (**A**) and post- (**B**) contrast images of NH showing isoattenuation and hypoenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and a cyst-like appearance. Pre- (**C**) and post- (**D**) contrast images of NH showing hypoattenuation and hypoenhancement, diffuse enhancement pattern, with heterogeneous distribution, ill-defined margins, regular surface, and cyst-like appearance. Pre- (**E**) and post- (**F**) contrast images of NH showing hypoattenuation and hyperenhancement, rim enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and solid appearance. The ROI is placed inside the lesions.

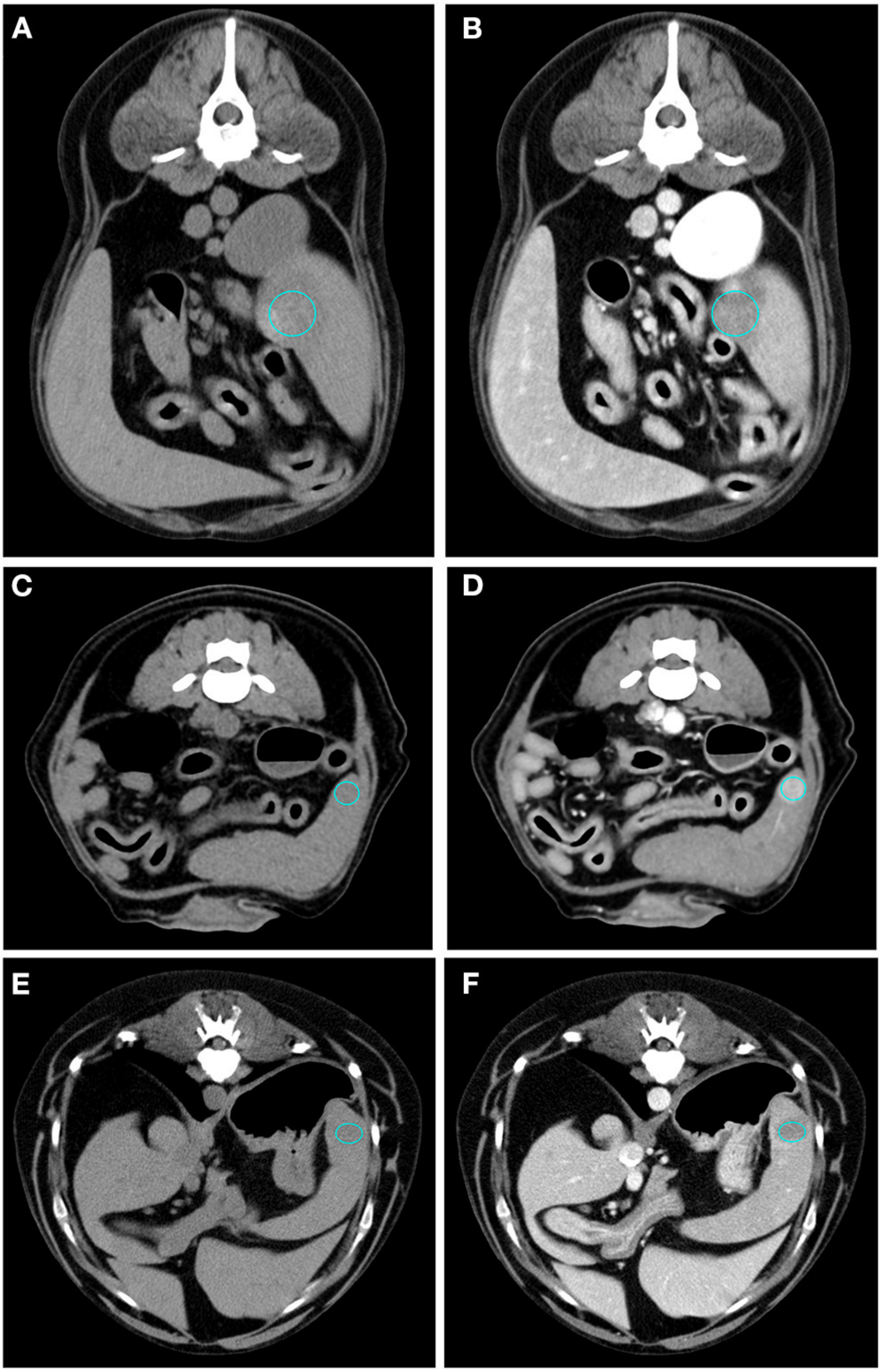


Figure 2 Pre- (**A**) and post- (**B**) contrast images of an OBL (diagnosed as extramedullary haematopoiesis) showing hypoattenuation and hypoenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and solid appearance. Pre- (**C**) and post- (**D**) contrast images of an OBL (diagnosed as extramedullary haematopoiesis) showing isoattenuation and hyperenhancement, diffuse enhancement pattern with homogeneous distribution, well-defined margins, regular surface, and solid appearance. Pre- (**E**) and post- (**F**) contrast images of an OBL (diagnosed as haematoma) showing hypoattenuation and hypoenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, regular surface, and cyst-like appearance. The ROI is placed inside the lesions.

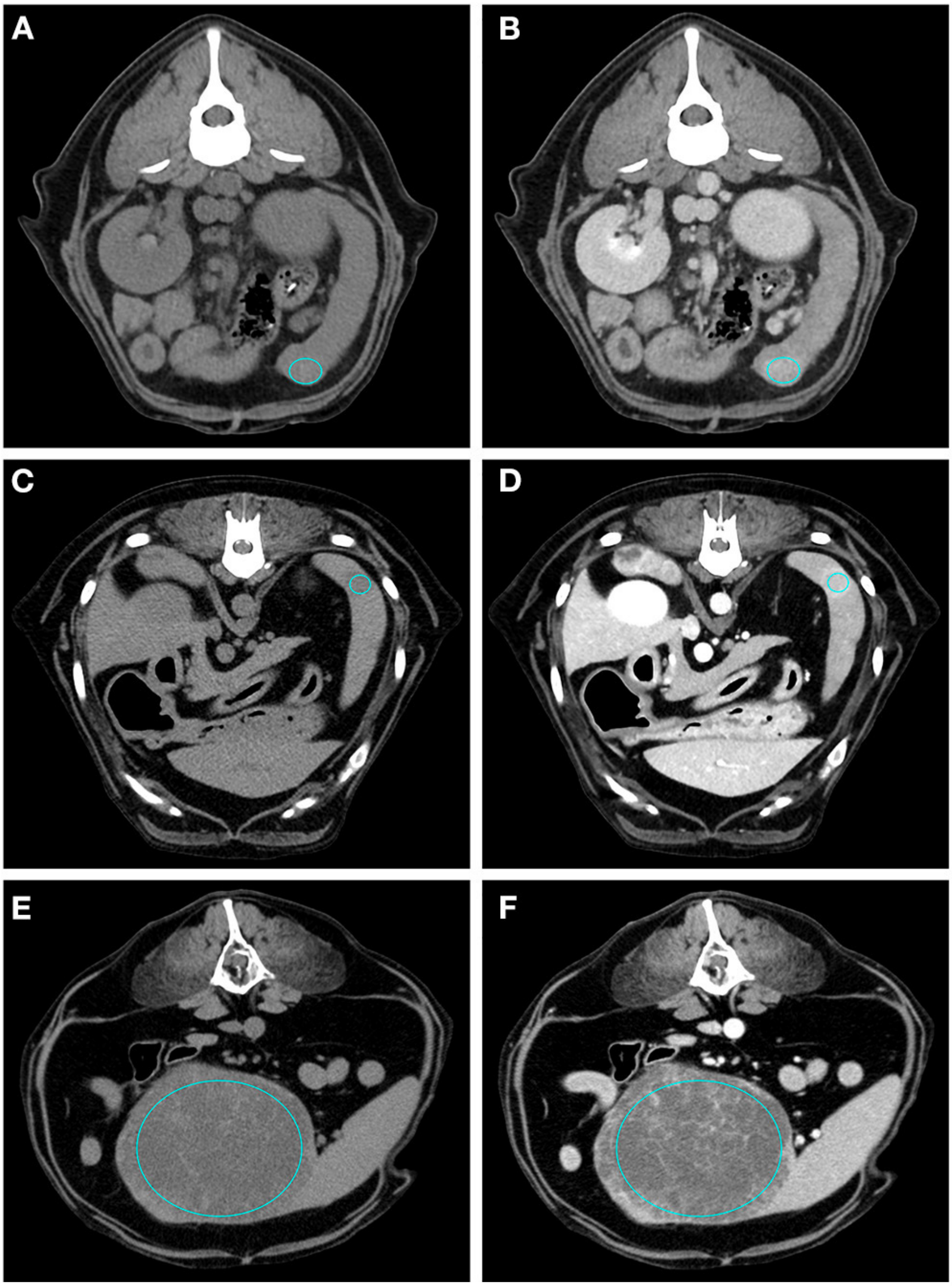


Figure 3 Pre- (**A**) and post- (**B**) contrast images of a RCT (diagnosed as lymphoma) showing isoattenuation and hyperenhancement, diffuse enhancement pattern with homogeneous distribution, ill-defined margins, irregular surface, and solid appearance. Pre- (**C**) and post- (**D**) contrast images of a RCT (diagnosed as mastocytoma) showing hypoattenuation and hyperenhancement, rim enhancement pattern, with heterogeneous distribution, well-defined margins, regular surface, and solid appearance. Pre- (**E**) and post- (**F**) contrast images of a RCT (diagnosed as mesenchymal neoplasia) showing hypoattenuation and hypoenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. The ROI is placed inside the lesions.

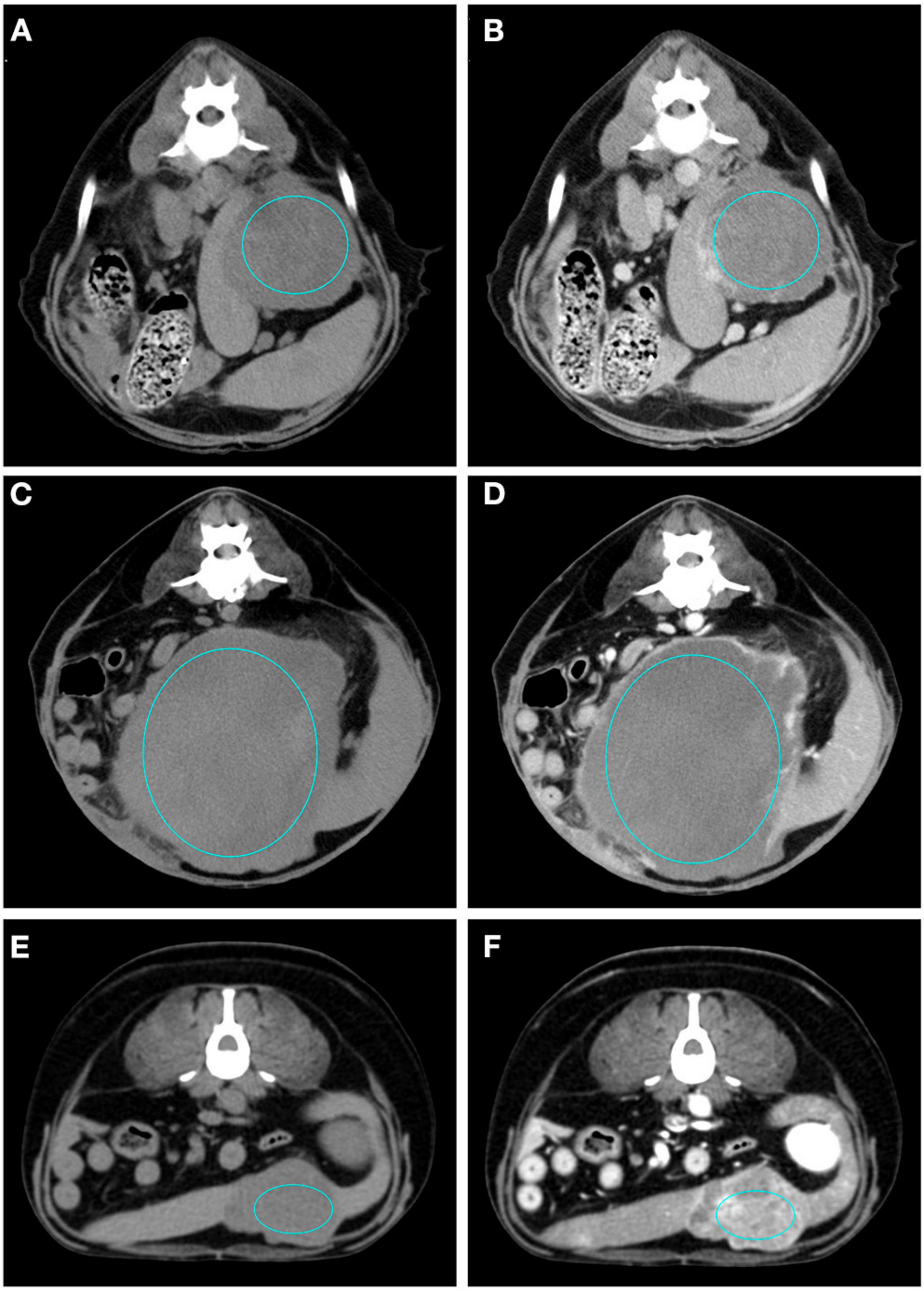


Figure 4 Pre- (**A**) and post- (**B**) contrast images of a sarcoma (diagnosed as haemangiosarcoma) showing hypoattenuation and hypoenhancement, rim enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. Pre- (**C**) and post- (**D**) contrast images of a sarcoma (diagnosed as sarcoma) showing isoattenuation and hypoenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. Pre- (**E**) and post- (**F**) contrast images of a sarcoma (diagnosed as stromal sarcoma) showing hypoattenuation and hyperenhancement, diffuse enhancement pattern with heterogeneous distribution, well-defined margins, irregular surface, and cyst-like appearance. The ROI is placed inside the lesions.

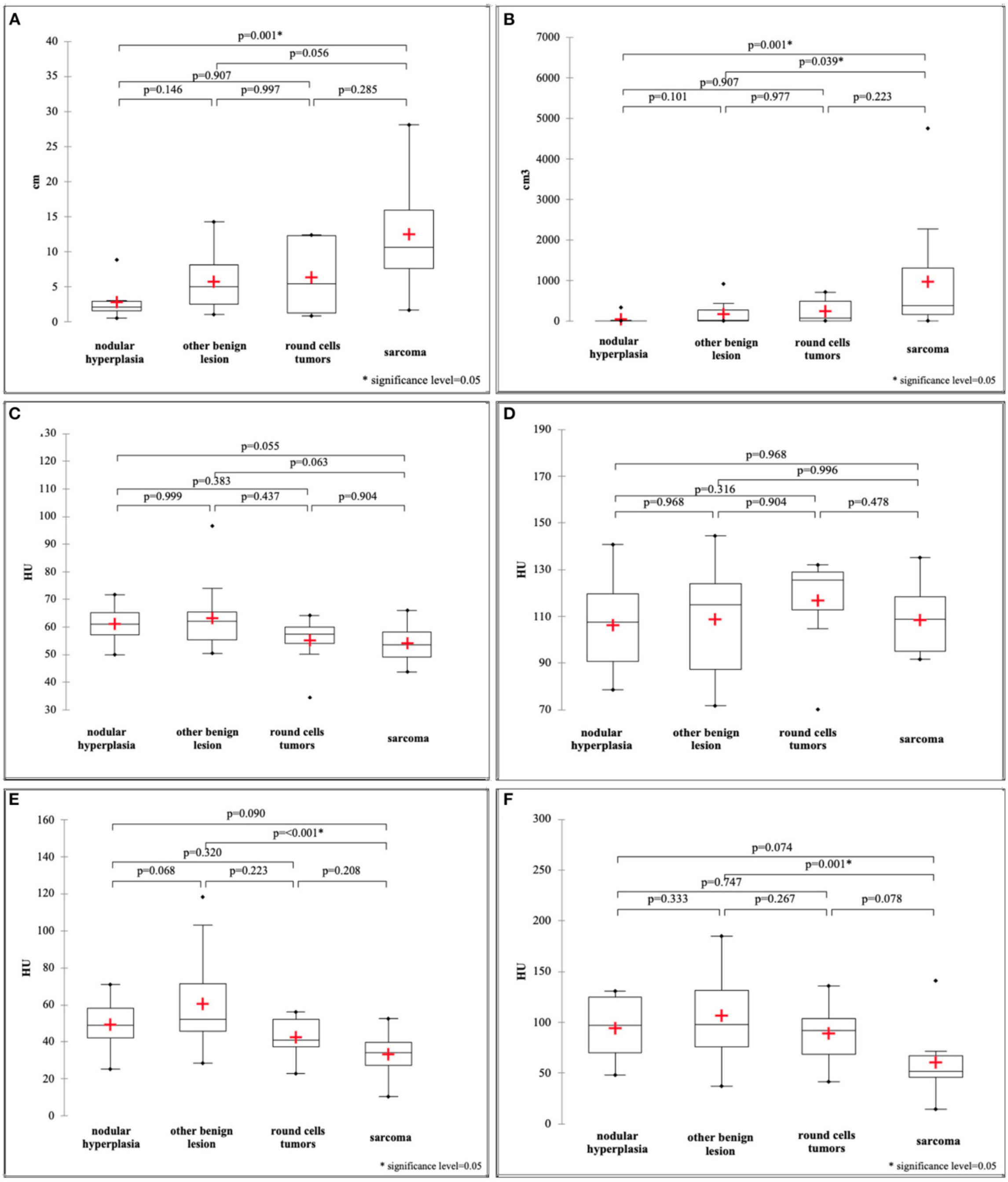


Figure 5 Box and whisker plot of the maximum dimension (**A**), ellipsoid volume (**B**), HU value of the pre-contrast normal spleen (**C**), HU value of the post-contrast normal spleen (**D**), HU value of the pre-contrast lesion (**E**), HU value of the post-contrast lesion (**F**).

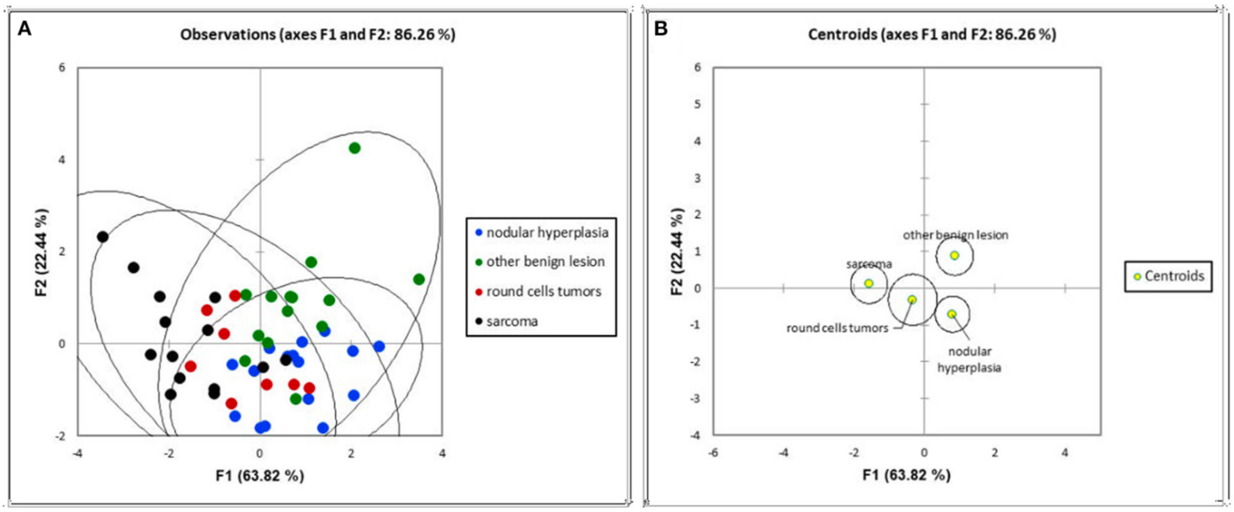


Figure 6 Distribution of the cases (A) and of the centroids (B) based on the F1 and F2 components of the factorial discriminant analysis classification.

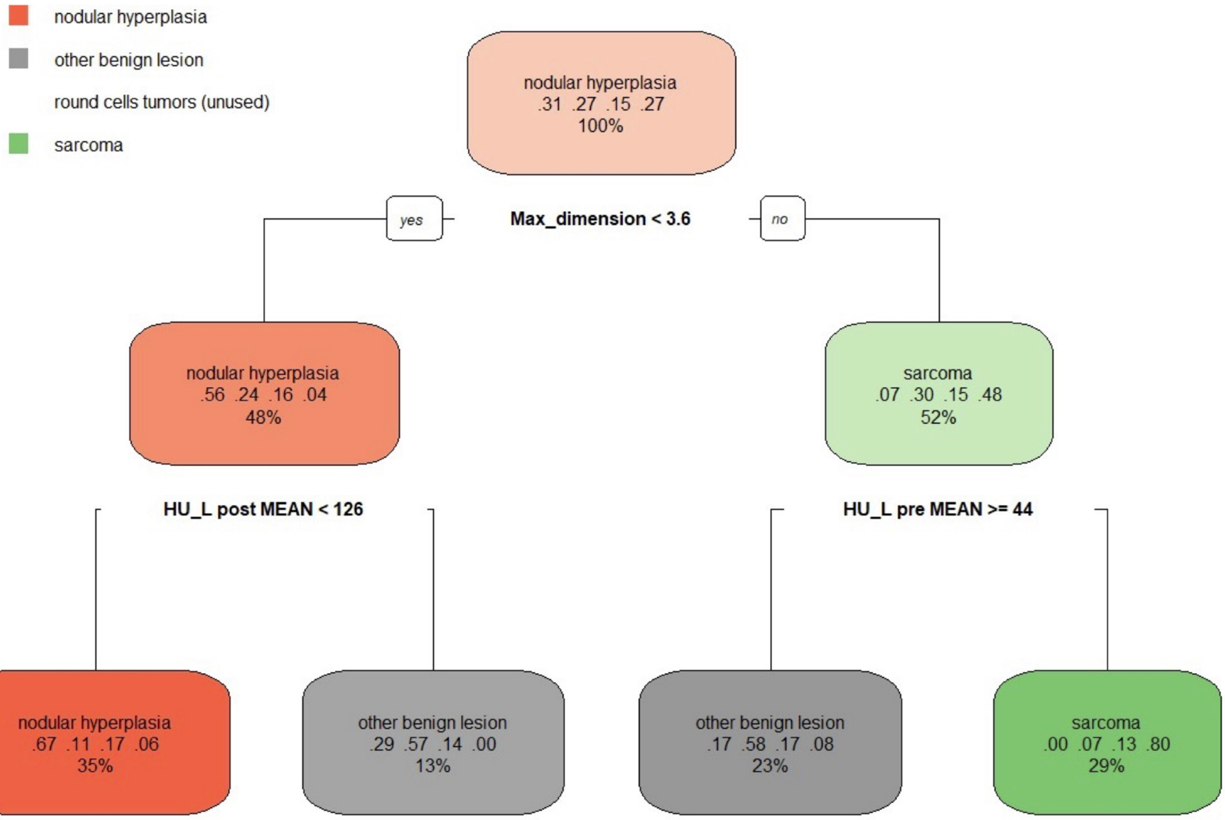


Figure 7 The machine learning-based decision tree developed on the qualitative and the quantitative CT features of the focal splenic lesions. The second line in each box shows the probability of each class at that node (i.e. the probability of the class conditioned on the node) and the third line shows the percentage of observations used at that node.

DISCUSSION

The complex relationship occurring between the CT features and the FSL histotypes were described using both a classical statistical and a machine learning-based approach. The classical statistical analysis revealed some significant differences between groups, mainly for the quantitative features, whereas no significant differences, with the exception of surface and appearance, were evident for the qualitative features. The machine learning algorithms substantially confirmed the results of the classical statistical analysis; indeed, all the features included in both in the FDA and the decision tree resulted as significantly different between the different groups in the classical statistical analysis tests. Nonetheless, the main advantage of FDA is that it allows identification of subtler trends than the classical statistical analysis does. As a result of the FDA, it emerged that while SAs were characterised both by larger dimensions and a cyst-like appearance, strong similarities were evident in the appearance of NH, OBLs and SAs than based on classical statistical analysis. Furthermore, the decision tree provided a straightforward and easy-to-use chart that could be directly used to classify lesions based on their CT features with a very high accuracy for SAs and a moderate accuracy for OBLs and NH, while RCTs could not be classified through the decision tree.

NH nodules are reported as having a variable appearance (both homogeneous and heterogeneous) and as being markedly hyperenhancing on post-contrast CT images (2,6,7). The results of both the discriminant analysis and the decision tree confirm such findings and, in fact, NHs were characterised by higher pre- and post-contrast HU values and a smaller volume than SAs were. Interestingly, but not unexpectedly since several extramedullary haematopoiesis cases were present in the OBL group, NH showed an overall lower enhancement than the OBLs did.

In this paper we placed all benign lesions other than NH (i.e. haematomas and extramedullary haematopoiesis) in the OBL group and, therefore, a straightforward comparison with the features reported by other authors for this lesion category is not possible. Haematomas are reported as mainly heterogeneously enhancing masses in all phases by both Kutara et al., 2017, and Jones et al., 2016. Splenic extramedullary haematopoiesis nodules have been described as having a very variable appearance and as hyperenhancing in all phases (4). The CT features of all the other possible OBLs (e.g.: splenitis, lipoma, etc.) have not yet been reported in the literature.

The literature reporting the CT features of malignant splenic lesions is fragmentary and different authors have grouped malignant lesions differently. Both Fife et al., 2004 and Jones et al., 2016, grouped all malignant lesions into a single category during statistical analysis. To date only Kutara et al., 2017 have considered sarcoma and haemangiosarcoma as individual categories during statistical analysis. In the present study, all the sarcoma cases were grouped together due to the presence of only 3 haemangiosarcomas in the database. To the best of the authors' knowledge, this is the first manuscript considering RCT as an individual category for the analysis. The results of the present study confirm the finding that sarcomas have a lower attenuation and larger dimensions compared to benign splenic lesions (both NH and OBLs). Instead, RCTs did not show any distinctive CT feature. Therefore, splenic lesions should not be classified based on their CT features alone since other diagnostic procedures (e.g.: cytology) are necessary to determine the histotypes. Interestingly, the splenic lymph nodes were normal in 63 % of RCTs and 43% of SAs, whereas the lymph nodes were abnormal in 23% of the benign lesions. Therefore, it is the author's opinion that lymph node evaluation also has poor value in determining whether a lesion is benign or malignant.

Fife et al., 2004, reported the presence of abdominal effusion as significantly correlated with the presence of malignant splenic lesions. Instead, abdominal effusion was not detected in any of the cases included in the present study and, therefore, the significance of such a finding was not evaluated. The significance of abdominal effusion was also not evaluated by Lee et al., 2017, and Jones et al., 2016.

Contrast-enhanced ultrasound has seldom been reported as useful in the differentiation between benign and malignant FSLs (24,25). However, these reports are largely outdated and based on a relatively low number of cases (26 and 29 cases respectively), and the full efficacy of this imaging technique has yet to be proven.

In human medicine, FSL diagnosis based on diagnostic imaging findings alone poses challenges similar to those of the veterinary context (26). Nonetheless, the combination of CT, MRI and 18F-FDG PET/CT enables attainment of a high degree of confidence for lesion characterisation (26). However, due to the limited availability of such imaging devices and of properly developed techniques (18F-FDG PET/CT) in veterinary medicine, the MRI features of FSLs have seldom been described in dogs (27). It is likely that the combination of different imaging techniques could enable attainment of a higher degree of confidence in FSL diagnosis also in the veterinary field. Nonetheless, this will require a broad standardisation of the features of FSLs in each individual diagnostic imaging technique.

Most of the results of the present study are in agreement with those reported in the literature. In particular, Lee et al. 2018 reported the HU values of the lesion in pre-contrast phase, along with the regular and irregular margination of the lesion as statistically significant features in the distinction between malignant and benign tumors. Results of this study confirmed, the appearance of the surface and the HU values of the lesion during pre-contrast phase to be useful in the distinction among the four considered pathological categories.

Fife et al. 2004 reported the HU values of the lesion during pre- and post-contrast phase as significantly different between malignant lesions, hematomas and NH. Nevertheless, the authors found a threshold of <55HU in post contrast scans for classification of malignant lesions. In our study, the decision tree, used a threshold of <44 HU to distinguish between sarcoma and OBL, and of <126 HU to distinguish between NH and OBL.

Lastly, Kutara et al. 2017 found the size of the mass as statistically different between NH, hematomas, hemangiosarcoma, and undifferentiated sarcoma. In particular, the size was smaller for NH. Our results confirmed that the size is significantly smaller in case of NH, with a 3.6 cm cut-off value.

One limitation of the present study is that, since most of the FSLs were occasional findings on CT scans performed for other reasons (e.g. staging of neoplasia), the lesions were evaluated only in the pre-contrast and in the delayed phase and no arterial phase was available for the selected cases. Previous reports on the CT features of FSL (2,6–8) used both dual-phase (6,7) and triple-phase (2,8) scanning protocols. To the best of the authors' knowledge, no CT features specifically related to the arterial phase have yet been shown as useful in differentiating between the different FSL histotypes.

Another limitation is that no specific RCTs were differentiated by the proposed decision tree. A possible explanation of this is that several different histotypes (lymphoma, histiocytic sarcoma, mastocytoma, mesenchymal neoplasia, plasma-cell neoplasia) with different imaging features were included in the RCT category. By including a larger number of cases in a future study, a larger number of groups could likely be considered during analysis, which would therefore provide a more detailed description of the CT features of FSLs.

The third possible limitation is related to the use of cytopathology to classify the cases. Indeed, the agreement between cytopathological and hystopathological diagnoses of the spleen is reported to be only moderate (Cohen's Kappa = 0.473) (28). In the present study histopathology was performed only in 19 cases, while the remainder 33 cases were evaluated only by means of cytopathology. To improve the classification accuracy both non-diagnostic cases and cases with doubtful cytological diagnosis were excluded. It is the authors' opinion that, including only cases with high quality cytopathological samples increase the diagnostic accuracy of the cytological exam.

CONCLUSION

The CT features of different groups of FSL have been described and analysed using both classical statistical analysis and machine learning algorithms. SAs are characterised by large dimensions, a cystic appearance and an overall low post-contrast enhancement. NH and OBLs are characterised by small dimensions, a solid appearance and a high post-contrast enhancement. OBLs show higher post-contrast values than NH. Lastly, RCTs do not exhibit any distinctive CT features. A straightforward, easy-to-use decision tree for classifying FSLs is proposed.

Funding The present paper is part of a project funded by a research grant from the Department of Animal Medicine, Production and Health – MAPS, University of Padua, Italy: SID - Zotti 2018 (€ 32,000; Application of deep-learning algorithms in pet animal diagnostic imaging)

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was conducted respecting Italian Legislative Decree N° 26/2014 (transposing EU directive 2010/63/EU). Since the data used in this study were part of routine clinical activity, no ethical committee approval was required. Informed consent for personal data processing was obtained from the owners.

Author contribution: TB, SB, and AZ conceived the study, performed the CT scans, and drafted and revised the manuscript. FB revised the manuscript and standardised part of the cytological examinations. BC drafted and revised the manuscript, and performed the statistical analysis. All the authors contributed to the article and approved the submitted version.

BIBLIOGRAPHY

1. Spangler WL, Culbertson MR. Prevalence, type, and importance of splenic diseases in dogs: 1,480 cases (1985-1989). *J Am Vet Med Assoc* (1992) **15**:829–834.
2. Kutara K, Seki M, Ishigaki K, Teshima K, Ishikawa C, Kagawa Y, Edamura K, Nakayama T, Asano K. Triple-phase helical computed tomography in dogs with solid splenic masses. *J Vet Med Sci* (2017) **79**:1870–1877. doi:10.1292/jvms.17-0253
3. Meuten DJ. *Tumors in Domestic Animals*. (2017). doi:10.1002/9780470376928
4. Cordella A, Caldin M, Bertolini G. Splenic extramedullary hematopoiesis in dogs is frequently detected on multiphase multidetector-row CT as hypervascular nodules. *Vet Radiol Ultrasound* (2020) **61**:512–518. doi:10.1111/vru.12872
5. Mattoon JSNTG. *SMALL ANIMAL DIAGNOSTIC ULTRASOUND*. (2015).
6. Fife WD, Samii VF, Drost T, Mattoon JS, Hoshaw-Woodard S. Comparison between malignant and nonmalignant splenic masses in dogs using contrast-enhanced computed tomography. *Vet Radiol Ultrasound* (2004) **45**:289–297. doi:10.1111/j.1740-8261.2004.04054.x
7. Jones ID, Lamb CR, Drees R, Priestnall SL, Mantis P. Associations between dual-phase computed tomography features and histopathologic diagnoses in 52 dogs with hepatic or splenic masses. *Vet Radiol Ultrasound* (2016) **57**:144–153. doi:10.1111/vru.12336
8. Lee M, Park J, Choi H, Lee H, Jeong SM. Presurgical assessment of splenic tumors in dogs: a retrospective study of 57 cases (2012–2017). *J Vet Sci* (2018) **19**:827. doi:10.4142/jvs.2018.19.6.827
9. Banzato T, Cherubini GB, Atzori M, Zotti A. Development of a deep convolutional neural network to predict grading of canine meningiomas from magnetic resonance images. *Vet J* (2018) **235**:90–92. doi:10.1016/j.tvjl.2018.04.001
10. Banzato T, Bernardini M, Cherubini GB, Zotti A. A methodological approach for deep learning to distinguish between meningiomas and gliomas on canine MR-images. *BMC Vet Res* (2018)

14:317. doi:10.1186/s12917-018-1638-2

11. Banzato T, Fiore E, Morgante M, Manuali E, Zotti A, Manu- E, Zotti A, Manuali E, Zotti A. Texture analysis of B-mode ultrasound images to stage hepatic lipidosis in the dairy cow: A methodological study. *Res Vet Sci* (2016) **108**:71–75. doi:10.1016/j.rvsc.2016.08.007
12. Banzato T, Bernardini M, Cherubini GB, Zotti A. Texture analysis of magnetic resonance images to predict histologic grade of meningiomas in dogs. *Am J Vet Res* (2017) **78**:1156–1162. doi:10.2460/ajvr.78.10.1156
13. Banzato T, Bonsembiante F, Aresu L, Gelain ME, Burti S, Zotti A. Use of transfer learning to detect diffuse degenerative hepatic diseases from ultrasound images in dogs: A methodological study. *Vet J* (2018) **233**:35–40. doi:10.1016/j.tvjl.2017.12.026
14. Li S, Wang Z, Visser LC, Wisner ER, Cheng H. Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. *Vet Radiol Ultrasound* (2020) **61**:611–618. doi:10.1111/vru.12901
15. McEvoy FJ, Proschowsky HF, Müller A V., Moorman L, Bender-Koch J, Svalastoga EL, Frellsen J, Nielsen DH. Deep transfer learning can be used for the detection of hip joints in pelvis radiographs and the classification of their hip dysplasia status. *Vet Radiol Ultrasound* (2021) **62**:387–393. doi:10.1111/vru.12968
16. Burti S, Osti VL, Zotti A, Banzato T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *Vet J* (2020) **262**:105505. doi:10.1016/j.tvjl.2020.105505
17. Banzato T, Wodzinski M, Burti S, Osti VL, Rossoni V, Atzori M, Zotti A. Automatic classification of canine thoracic radiographs using deep learning. *Sci Rep* (2021) **11**:1–8. doi:10.1038/s41598-021-83515-3
18. Banzato T, Causin F, Puppa A Della, Cester G, Mazzai L, Zotti A, Della Puppa A, Cester G, Mazzai L, Zotti A. Accuracy of Deep Learning to Differentiate the Histopathological Grading of

Meningiomas on MR Images: A Preliminary Study. *J Magn Reson Imaging* (2019)1–8. doi:10.1002/jmri.26723

19. Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. *Front Vet Sci* (2021) **8**:1–10. doi:10.3389/fvets.2021.611556
20. Fukushima K, Kanemoto H, Ohno K, Takahashi M, Nakashima K, Fujino Y, Uchida K, Fujiwara R, Nishimura R, Tsujimoto H. CT CHARACTERISTICS OF PRIMARY HEPATIC MASS LESIONS IN DOGS. *Vet Radiol Ultrasound* (2012) **53**:n/a-n/a. doi:10.1111/j.1740-8261.2011.01917.x
21. Sápi J, Kovács L, Drexler DA, Kocsis P, Gajári D, Sápi Z. Tumor volume estimation and quasi-continuous administration for most effective bevacizumab therapy. *PLoS One* (2015) **10**:1–20. doi:10.1371/journal.pone.0142190
22. Kutara K, Seki M, Ishikawa C, Sakai M, Kagawa Y, Iida G, Ishigaki K, Teshima K, Edamura K, Nakayama T, et al. TRIPLE-PHASE HELICAL COMPUTED TOMOGRAPHY IN DOGS WITH HEPATIC MASSES. *Vet Radiol Ultrasound* (2014) **55**:7–15. doi:10.1111/vru.12099
23. Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. *Front Vet Sci* (2021) **8**:1–10. doi:10.3389/fvets.2021.611556
24. Mangano C, Macrì F, Di Pietro S, Pugliese M, Santoro S, Iannelli NM, Mazzullo G, Crupi R, De Majo M. Use of contrast-enhanced ultrasound for assessment of nodular lymphoid hyperplasia (NLH) in canine spleen. *BMC Vet Res* (2019) **15**:1–9. doi:10.1186/s12917-019-1942-5
25. Rossi F, Leone VF, Vignoli M, Laddaga E, Terragni R. Use of contrast-enhanced ultrasound for characterization of focal splenic lesions. *Vet Radiol Ultrasound* (2008) **49**:154–164. doi:10.1111/j.1740-8261.2008.00343.x
26. Barat M, Hoeffel C, Aissaoui M, Dohan A, Oudjit A, Dautry R, Paisant A, Malgras B, Cottereau

AS, Soyer P. Focal splenic lesions: Imaging spectrum of diseases on CT, MRI and PET/CT. *Diagn Interv Imaging* (2021) **102**:501–513. doi:10.1016/j.diii.2021.03.006

27. Clifford CA, Pretorius ES, Weisse C, Sorenmo KU, Drobatz KJ, Siegelman ES, Solomon JA. Magnetic Resonance Imaging of Focal Splenic and Hepatic Lesions in the Dog. *J Vet Intern Med* (2004) **18**:330–338. doi:10.1111/j.1939-1676.2004.tb02554.x

28. Tecilla M, Gambini M, Forlani A, Caniatti M, Ghisleni G, Roccabianca P. Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. *PLoS One* (2019) **14**:1–15. doi:10.1371/journal.pone.0224945

CHAPTER FIVE

Application of a convolutional neural network to CT images of focal liver and splenic lesions.

IMAGE PREPROCESSING

Starting from the same patients included in the descriptive study (69 cases of liver lesions, 52 cases of splenic lesions), we have exported all the images in a Digital Imaging and Communications in Medicine (DICOM) format into two folders, namely training set and test set, both containing two subfolders: one for benign, and one for malignant cases. 15% of the patients was included in the test set, whereas the remaining 85% was included in the training set. Therefore, the hepatic lesions test set was composed of the images from 5 benign and 5 malignant cases; the splenic lesions test set was composed of the images from 4 benign and 4 malignant cases. All the examination were reviewed with Osirix 10.0.4 program. Using the drawing tool, the masks were manually drowned, following the contour of each lesion in every slice where was visible. The pixel values inside the masks were set to 1, and the pixel values outside the masks were set to 0 (Figure 1 and 2). In this way, we have obtained a spatial representation of lesion inside the patient. The masks were exported, and two volumes were created using 3D-Slicer (v. 4.10.1) [1]: 1) a binary volume containing the information of the mask; and 2) a volume containing the reconstruction of the original tomographic examination. The aim of the mask was to allow the algorithm to identify the coordinates for the localization of the lesion in the patient and then to extract the lesion from the volumes. One of the big challenges was to identify the best images cropping method. Neural networks are only capable of processing square images with constant dimensions (i.e. 224 x 224 pixel) that are specific to the algorithm used. As a consequence, if we use the dimension of the lesion as crop factor, the smaller lesions will be magnificated leading to a loss of dimensional information. On the other hand, the biggest lesions will be compressed to smaller dimensions with strong geometrical distortions of the image. Instead, if we decide to crop the lesions using a fixed dimension (i.e. 120 x 120 pixel), some lesions will be very small compared to the included area, leading to a difficult identification by the deep-learning algorithm. The perfect approach does not exist, so we opted for a compromise: using an on-purpose developed Python (v. 3.9.0) script, we matched the two volumes and the lesions were cropped with a 0.6 factor, based on the dimension of the biggest lesion. In this way, the dimensions of the smaller lesions were preserved and only a few portions of the largest lesions were excluded from the crop. Finally, only the slices containing at least 200 pixels of the lesion were sent to the CNN.

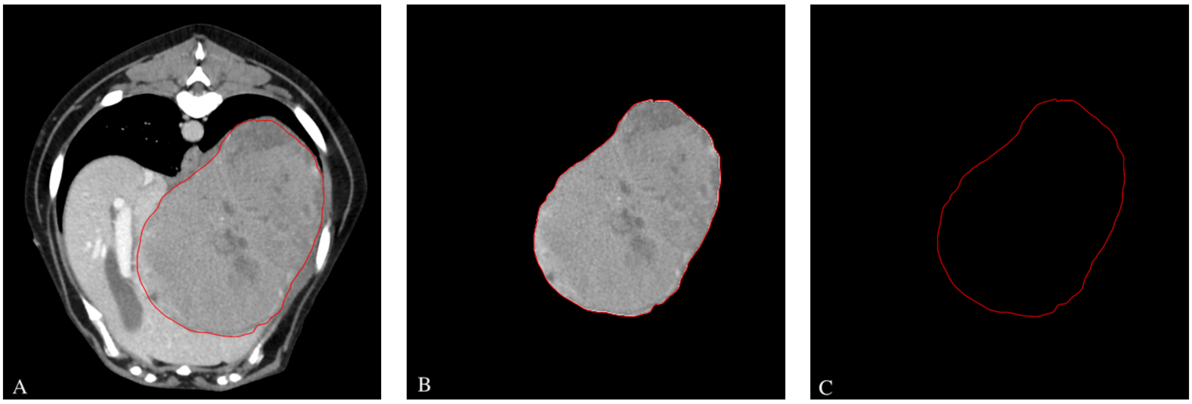


Figure 1 Example of a mask drowned around the contour of a hepatocarcinoma on a single CT slice. A. The mask is drowned B. Pixel volumes external to the mask is changed to 0 C. Pixel volumes internal to the mask is changed to 1.

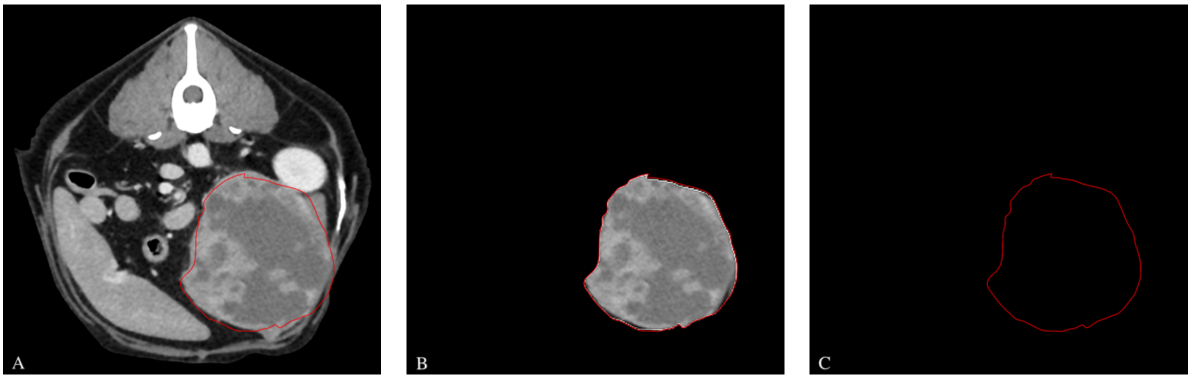


Figure 2 Example of a mask drowned around the contour of a splenic sarcoma on a single CT slice. A. The mask is drowned B. Pixel volumes external to the mask is changed to 0 C. Pixel volumes internal to the mask is changed to 1.

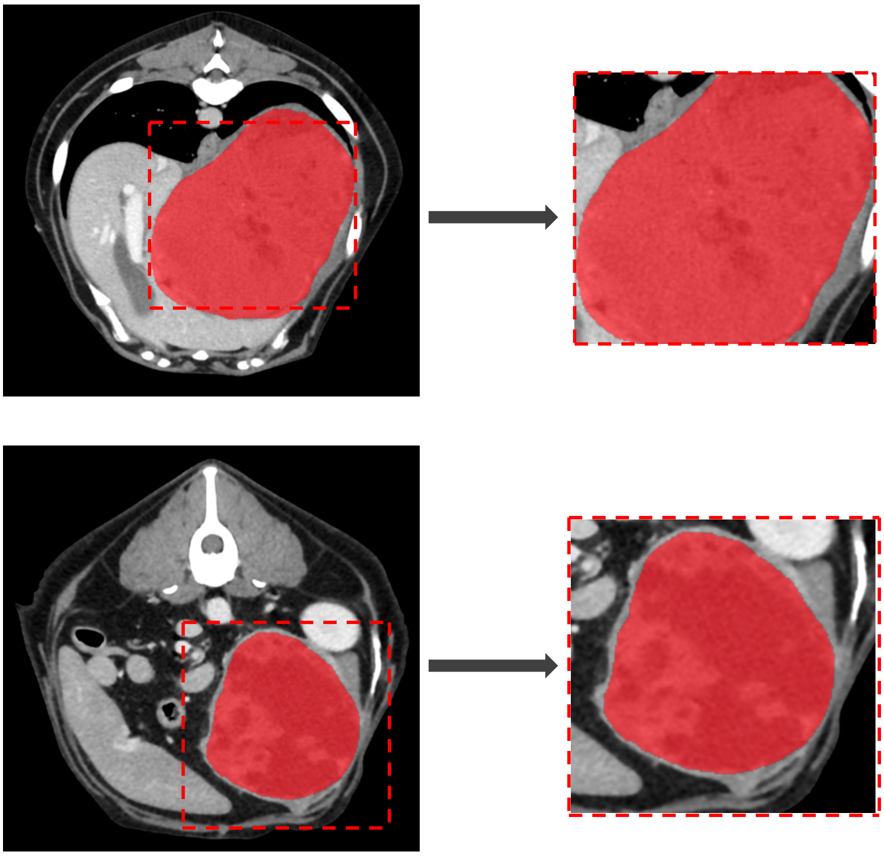


Figure 3 Example of crop of a lesion with crop factor of 0.6. In the upper part a hepatocarcinoma is reported; in the bottom part a splenic sarcoma is reported.

DEEP LEARNING

For the deep-learning analysis a dedicated work-station (Linux operating system, Ubuntu 18.04, Canonical) with 4 GPU units (Tesla V100; NVIDIA), 2.2 GHz processor (Intel Xeon E5-2698 v4; Intel), and 256 of memory RAM was used. The pre-trained on ImageNet Incetion V3 [2,3] CNN was used.

One of the limitation of the use of neural networks on small datasets (like in our study) is represented by overfitting, also known as overtraining; when a model is trained on a small dataset, it might be able to learn the typical features describing the test dataset, but then has a poor generalization ability on new datasets [4]. In other words, the model will perform well in the training set, but its performances will be significantly lower on new images. Different methods were proposed to overcome (at least partially) the overfitting problem. Among others there is data augmentation. With this method new synthetic samples are created from the original dataset. These new samples are added to the original dataset resulting in an augmentation of its dimension and, therefore, in the reduction of overfitting. As a consequence, the model will have a better generalization ability on new datasets. In our study, we applied the following data augmentation methods: cropping (the images were cropped in sub-images starting from each slice of the CT examination); translation (images are translated along the x and y axes); flipping (images are mirrored horizontally and vertically); rotation (images are rotated clock-wise and counterclock-wise); changes in the contrast levels of the images.

CASES CLASSIFICATION

The Inception V3 performances were evaluated by means of accuracy, sensitivity, and specificity, considering a p-value of 0.05. Since every case was composed by hundreds of images and the algorithm classifies each image independently, every case of the test set was classified as malignant if at least the 60% of the images was classified as malignant.

Using this scheme, Inception V3 correctly classified 7 out of 10 cases of the hepatic lesions (4 malignant and 3 benign), and 5 out of 8 cases of the splenic lesions (3 malignant and 2 benign) in the test set. The confusion matrix of the hepatic and splenic lesions classification results is reported in Table 1 and 2 respectively. Accuracy, sensitivity, and specificity results are reported in Table 3.

Table 4 Confusion matrix of the classification of liver lesions with Inception V3.

		Reale		
		Maligno	**Benigno**	**Totale**
Predetto	**Maligno**	4	2	6
	Benigno	1	3	4
	Totale	5	5	10

Table 2 Confusion matrix of the classification of splenic lesions with Inception V3.

		Reale		
		Maligno	**Benigno**	**Totale**
Predetto	**Maligno**	3	2	5
	Benigno	1	2	3
	Totale	4	4	8

Table 5 Performance of Inception V3 in the classification od liver and splenic lesions.

	Accuratezza	**Sensibilità**	**Specificità**	**PLR**	**NLR**
Masse epatiche	0.70	0.80	0.60	2.00	3.00
Masse spleniche	0.63	0.75	0.50	1.50	2.00

CONCLUSIONS

To the best of our knowledge, this is the first study exploring the application of CNNs to distinguish between malignant and benign canine focal hepatic and splenic lesions based on CT studies. This distinction is crucial, since the treatment and the prognosis of the patient are closely related to the nature of the lesion. At the present time the gold standard diagnostic technique for the characterization of hepatic and splenic masses remains histopathology, despite being invasive and risky for the patient [5–7].

In human medicine, commercial products applying deep learning in the diagnostic imaging field are already available [8–10], and are considered as teh state-of-the-art to solve common tasks like lesion identification, segmentation, and classification [11]. Indeed, deep learning algorithms have demonstrates high level of accuracy (96-100%) in the identification of focal liver lesions and steatosis, and of 90% in the discrimination among the different tumoral histotype on the ultrasonographic images [9]. In the study of Zhou et al 2019 they highlight that the deep-learning algorithms are able to identify liver tumoral masses and metastatic lesions with 86% and 90% of accuracy respectively, when applied to CT images. Also in human medicine no advanced studies exploring the applications of deep learning on focal splenic lesions are available.

In the veterinary literature, a study [12] proposing a machine-learning based model applied to hepatic lesions showed an accuracy of 0.90 (sensitivity 0.67, specificity 1.00) in the distinction between malignant and benign lesions. In their model the features of voxel volume and uniformity resulted to be more informative for the distinction. Nevertheless, in their study almost only quantitative features were considered (with only 2 qualitative features). On the other hand, they have used a machine-learning model that usually perform better than deep learning on small datasets. By the way, in this type of model the human intervention is necessary to decide the features to be considered for the classification. In our study, both quantitative and qualitative features were considered and the CNN automatically extracted the features for the classification, without no human intervention.

The CNN Inception V3 correctly classified the 70% of the cases of the hepatic test set, with a good level of accuracy (0.80) and sensitivity (0.80), but low specificity (0.60). This means that the model has a good accuracy in the identification the malignant liver lesions, but a low accuracy in the distinction of benign lesions. On the other hand, the CNN developed on splenic lesions showed good levels of sensitivity (0.75), but a low specificity (0.50) and accuracy (0.63), correctly classifying only the 62.5% of the cases. It appears evident that one the main limitations of this study was the limited dimensions of the dataset. Indeed, the high performance that these algorithms reach in human medicine[11,13,14] arise from a very large training set containing hundreds of thousand images. The results of this study are promising but still very in a very embryonal phase, a first step in the future, through completely new horizons for the veterinary medicine. One of the main advantages of deep learning algorithms is that both they can be easily implemented by increasing the dimensions of the training dataset. Furthermore, there is a very active research aimed to develop new and more preforming deep learning algorithms specifically designed for medical images. This project is not concluded: more cases will be added in the future.

BIBLIOGRAFIA

1. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30(9):1323-1341. doi:10.1016/j.mri.2012.05.001
2. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016-Decem:2818-2826. doi:10.1109/CVPR.2016.308
3. Szegedy C, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *Proc Thirty-First AAAI Conf Artif Intell San Fr CA, USA*.:4278-4284. http://arxiv.org/abs/1512.00567.
4. Sahiner B, Pezeshk A, Hadjiiski LM, et al. Deep learning in medical imaging and radiation therapy. *Med Phys*. 2019;46(1):e1-e36. doi:10.1002/mp.13264
5. McDevitt HL, Mayhew PD, Giuffrida MA, Brown DC, Culp WTN, Runge JJ. Short-term clinical outcome of laparoscopic liver biopsy in dogs: 106 cases (2003-2013). *J Am Vet Med Assoc*. 2016;248(1):83-90. doi:10.2460/javma.248.1.83
6. Reece J, Pavlick M, Penninck DG, Webster CRL. Hemorrhage and complications associated with percutaneous ultrasound guided liver biopsy in dogs. *J Vet Intern Med*. 2020;34(6):2398-2404. doi:10.1111/jvim.15942
7. Tecilla M, Gambini M, Forlani A, Caniatti M, Ghisleni G, Roccabianca P. Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. *PLoS One*. 2019;14(11):1-15. doi:10.1371/journal.pone.0224945
8. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. *PLoS Med*. 2018;15(11):1-15. doi:10.1371/journal.pmed.1002697
9. Zhou LQ, Wang JY, Yu SY, et al. Artificial intelligence in medical imaging of the liver. *World J Gastroenterol*. 2019;25(6):672-682. doi:10.3748/wjg.v25.i6.672
10. Che H, Brown LG, Foran DJ, Nosher JL, Hacihaliloglu I. Liver disease classification from ultrasound using multi-scale CNN. *Int J Comput Assist Radiol Surg*. 2021;16(9):1537-1548. doi:10.1007/s11548-021-02414-0
11. Wang L, Wang H, Huang Y, et al. Trends in the application of deep learning networks in medical image analysis: Evolution between 2012 and 2020. *Eur J Radiol*. 2022;146:110069. doi:10.1016/j.ejrad.2021.110069

12. Shaker R, Wilke C, Ober C, Lawrence J. Machine learning model development for quantitative analysis of CT heterogeneity in canine hepatic masses may predict histologic malignancy. *Vet Radiol Ultrasound*. 2021;62(6):711-719. doi:10.1111/vru.13012
13. Moawad AW, Fuentes DT, Elbanan MG, et al. Artificial Intelligence in Diagnostic Radiology: Where Do We Stand, Challenges, and Opportunities. *J Comput Assist Tomogr*. 2022;46(1):78-90. doi:10.1097/RCT.0000000000001247
14. Biswas M, Kuppili V, Saba L, et al. State-of-the-art review on deep learning in medical imaging. *Front Biosci - Landmark*. 2019;24(3):392-426. doi:10.2741/4725

CHAPTER SIX

Other research projects

A FRAILTY INDEX BASED ON CLINICAL DATA TO QUANTIFY MORTALITY RISK IN DOGS

Banzato T, Franzo G, Di Maggio R, Nicoletto E, Burti S, Cesari M, Canevelli M. **A Frailty Index based on clinical data to quantify mortality risk in dogs.** Sci Rep. 2019 Nov 14;9(1):16749. doi: 10.1038/s41598-019-52585-9. PMID: 31727920; PMCID: PMC6856105 [1]

Frailty is defined as a decline in an organism's physiological reserves resulting in increased vulnerability to stressors. In humans, a single continuous variable, the so-called Frailty Index (FI), can be obtained by multidimensionally assessing the biological complexity of an ageing organism. Here, we evaluate this variability in dogs and compare it to the data available for humans. In dogs, there was a moderate correlation between age and the FI, and the distribution of the FI increased with age. Deficit accumulation was strongly related to mortality. The effect of age, when combined with the FI, was negligible. No sex-related differences were evident. The FI could be considered in epidemiological studies and/or experimental trials to account for the potential confounding effects of the health status of individual dogs. The age-related deficit accumulation reported in dogs is similar to that demonstrated in humans. Therefore, dogs might represent an excellent model for human aging studies.

CONTRAST-ENHANCED ULTRASONOGRAPHY FEATURES OF HEPATOBILIARY NEOPLASMS IN CATS

Banzato T, Burti S, Rubini G, Orlandi R, Bargellini P, Bonsembiante F, Zotti A. **Contrast-enhanced ultrasonography features of hepatobiliary neoplasms in cats.** Vet Rec. 2020 Mar 14;186(10):320. doi: 10.1136/vr.105453. Epub 2019 Oct 3. PMID: 31582574; PMCID: PMC7079193. [2]

Background Contrast-enhanced ultrasonography (CEUS) features of primary hepatobiliary neoplasms have been reported in dogs but no information is available in cats.

Methods Qualitative and quantitative features of bile duct adenomas (BDAs, n=20), bile duct carcinomas (BDCs, n=16), and hepatocellular carcinomas (HCCs, n=8) are described in 44 cats.

Results There was an overlap in CEUS qualitative features between different histotypes, both in wash-in and wash-out phases. Distinction between different neoplasms based only on the CEUS qualitative features was not possible. At peak of enhancement, the BDAs, BDCs and HCCs showed a large range of echogenicities, from hypoenhancement to hyperenhancement, in comparison to the liver parenchyma. Eight of 20 BDAs showed inhomogeneous hyperenhancement during wash-in, which is a feature reported as typical of malignant lesions in dogs. BDC had a significantly faster wash-in compared with both BDA and HCC but the diagnostic accuracy of all the included quantitative variables was only moderate. No significant differences in the wash-out quantitative features of BDA and BDC were evident.

Conclusion There is poor evidence that CEUS may be used to distinguish between different primary hepatobiliary neoplasms in cats.

CORRELATION BETWEEN RENAL HISTOPATHOLOGY AND RENAL ULTRASOUND IN DOGS

Burti S, Zotti A, Bonsembiante F, Mastellaro G, Banzato T. **Correlation between renal histopathology and renal ultrasound in dogs.** Res Vet Sci. 2020 Apr;129:59-65. doi: 10.1016/j.rvsc.2020.01.003. Epub 2020 Jan 3. PMID: 31931264. [3]

Fifty-three privately owned dogs were included in the study. Ultrasonography of the kidneys was performed ante mortem. All the dogs died or were euthanized for reasons unrelated to this study. Histopathology of both kidneys was performed, and a degeneration and an inflammation score ranging from zero to two was assigned by consensus between two pathologists. A numerical score based on a three level semi-quantitative scale (0, 0.5, 1) was assigned by consensus between two of the authors to the following ultrasonographic abnormalities: corticomedullary definition, echogenicity of the renal cortex, echogenicity of the medulla, renal shape, cysts, scars, mineralizations, subcapsular perirenal fluid accumulation, pyelectasia. The scores deriving from the consensus were summed to create a summary index called renal ultrasound score (RUS). Statistically significant differences in cortico-medullary definition, echogenicity of the renal cortex, echogenicity of the medulla, renal shape, scars and pyelectasia were evident between the degeneration score groups. There were significantly different distributions of cortico-medullary definition, renal shape and scars between the inflammatory score groups. There were statistically significant differences in the RUS between the degenerative score groups (F = 24.154, p-value<.001). Post-hoc tests revealed significant differences between all groups. There were no significant differences in the RUS between the inflammatory score groups (F = 1.312, p-value = .264). Post-hoc tests revealed no significant differences between groups. The results of the present study suggest that the number and severity of the ultrasonographic abnormalities are correlated with the severity of the kidney degeneration. On the other hand, inflammation showed poor influence on the ultrasonographic appearance of the kidneys.

CONTRAST-ENHANCED ULTRASOUND FEATURES OF MALIGNANT FOCAL LIVER MASSES IN DOGS

Burti S, Zotti A, Rubini G, Orlandi R, Bargellini P, Bonsembiante F, Banzato T. **Contrast-enhanced ultrasound features of malignant focal liver masses in dogs.** Sci Rep. 2020 Apr 8;10(1):6076. doi: 10.1038/s41598-020-63220-3. PMID: 32269300; PMCID: PMC7142119. [4]

A total of 185 cases (150 retrospectively and 35 prospectively) of malignant liver masses were collected. In the retrospectively collected cases hyperenhancement during wash-in was the most common feature in HCCs but there was a high percentage of cases showing no enhancement or hypo/isoenhancement. ICCs displayed a large variety of contrast enhancement patterns and, although statically significant differences between ICCs and HCCs were evident, no clear distinction between these two pathologies was possible based only on their CEUS appearance. Sarcomas displayed all the possible degrees of wash-in enhancement with non-enhancing being the most common appearance. Metastases displayed all the possible contrast-enhancement patterns, with the most common being hyperenhancement in the wash-in phase followed by hypoenhancement in the wash-out phase. A decision tree was developed based on the features of the retrospectively selected cases. Based on the developed decision tree 27/35 prospectively collected cases were correctly classified. Even if some significant differences among groups were evident, all the histotypes displayed all the possible patterns of contrast enhancement, and, therefore, the differentiation of liver masses in dogs based only on their CEUS features is not feasible and, therefore, cytology or histopathology is required.

USE OF DEEP LEARNING TO DETECT CARDIOMEGALY ON THORACIC RADIOGRAPHS IN DOGS

Burti S, Longhin Osti V, Zotti A, Banzato T. **Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs**. Vet J. 2020 Aug;262:105505. doi: 10.1016/j.tvjl.2020.105505. Epub 2020 Jul 7. PMID: 32792095. [5]

The purpose of this study was to develop a computer-aided detection (CAD) device based on convolutional neural networks (CNNs) to detect cardiomegaly from plain radiographs in dogs. Right lateral chest radiographs (n = 1465) were retrospectively selected from archives. The radiographs were classified as having a normal cardiac silhouette (No-vertebral heart scale [VHS]-Cardiomegaly) or an enlarged cardiac silhouette (VHS-Cardiomegaly) based on the breed-specific VHS. The database was divided into a training set (1153 images) and a test set (315 images). The diagnostic accuracy of four different CNN models in the detection of cardiomegaly was calculated using the test set. All tested models had an area under the curve >0.9, demonstrating high diagnostic accuracy. There was a statistically significant difference between Model C and the remainder models (Model A vs. Model C, P = 0.0298; Model B vs. Model C, P = 0.003; Model C vs. Model D, P = 0.0018), but there were no significant differences between other combinations of models (Model A vs. Model B, P = 0.395; Model A vs. Model D, P = 0.128; Model B vs. Model D, P = 0.373). Convolutional neural networks could therefore assist veterinarians in detecting cardiomegaly in dogs from plain radiographs.

AUTOMATIC CLASSIFICATION OF CANINE THORACIC RADIOGRAPHS USING DEEP LEARNING

Banzato T, Wodzinski M, Burti S, Osti VL, Rossoni V, Atzori M, Zotti A. **Automatic classification of canine thoracic radiographs using deep learning.** Sci Rep. 2021 Feb 17;11(1):3964. doi: 10.1038/s41598-021-83515-3. PMID: 33597566; PMCID: PMC7889925. [6]

The interpretation of thoracic radiographs is a challenging and error-prone task for veterinarians. Despite recent advancements in machine learning and computer vision, the development of computer-aided diagnostic systems for radiographs remains a challenging and unsolved problem, particularly in the context of veterinary medicine. In this study, a novel method, based on multilabel deep convolutional neural network (CNN), for the classification of thoracic radiographs in dogs was developed. All the thoracic radiographs of dogs performed between 2010 and 2020 in the institution were retrospectively collected. Radiographs were taken with two different radiograph acquisition systems and were divided into two data sets accordingly. One data set (Data Set 1) was used for training and testing and another data set (Data Set 2) was used to test the generalization ability of the CNNs. Radiographic findings used as non-mutually exclusive labels to train the CNNs were: unremarkable, cardiomegaly, alveolar pattern, bronchial pattern, interstitial pattern, mass, pleural effusion, pneumothorax, and megaesophagus. Two different CNNs, based on ResNet-50 and DenseNet-121 architectures respectively, were developed and tested. The CNN based on ResNet-50 had an Area Under the Receive-Operator Curve (AUC) above 0.8 for all the included radiographic findings except for bronchial and interstitial patterns both on Data Set 1 and Data Set 2. The CNN based on DenseNet-121 had a lower overall performance. Statistically significant differences in the generalization ability between the two CNNs were evident, with the CNN based on ResNet-50 showing better performance for alveolar pattern, interstitial pattern, megaesophagus, and pneumothorax.

BIBLIOGRAPHY

1. Banzato T, Franzo G, Di Maggio R, et al. A Frailty Index based on clinical data to quantify mortality risk in dogs. *Sci Rep*. 2019;9(1):2-10. doi:10.1038/s41598-019-52585-9
2. Banzato T, Burti S, Rubini G, et al. Contrast-enhanced ultrasonography features of hepatobiliary neoplasms in cats. *Vet Rec*. 2019:vetrec-2019-105453. doi:10.1136/vr.105453
3. Burti S, Zotti A, Bonsembiante F, Mastellaro G, Banzato T. Correlation between renal histopathology and renal ultrasound in dogs. *Res Vet Sci*. 2020;129(January):59-65. doi:10.1016/j.rvsc.2020.01.003
4. Burti S, Zotti A, Rubini G, et al. Contrast-enhanced ultrasound features of malignant focal liver masses in dogs. *Sci Rep*. 2020;10(1):1-12. doi:10.1038/s41598-020-63220-3
5. Burti S, Longhin Osti V, Zotti A, Banzato T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *Vet J*. 2020;262:105505. doi:10.1016/j.tvjl.2020.105505
6. Banzato T, Wodzinski M, Burti S, et al. Automatic classification of canine thoracic radiographs using deep learning. *Sci Rep*. 2021;11(1):1-8. doi:10.1038/s41598-021-83515-3

CHAPTER SEVEN

Conclusions

The present study was structured in two main steps. In the first step I have applied some machine learning algorithms (decision tree mainly) to distinguish among different focal hepatic and splenic lesions based on their tomographic features. Furthermore, a metanalysis study to characterise the accuracy of individual CT features of focal liver lesions in the distinction between benign and malignant lesions was conducted. In the second step of my project a deep-learning convolutional neural network to classify malignant and benign focal liver and splenic lesions based on the CT images was developed.

Based on the results of my study HCCs are usually characterised by heterogeneous and cyst-like appearance, and hypoattenuating in the delayed phase. These results are only partially in agreement with the previous literature: indeed, all the possible enhancement patterns in post-contrast phases are described for HCC[1–8]. On the other hand, all the enhancement patterns nodular hyperplasia is reported in fact also in these lesions are possible[1–8].

The proposed machine-learning based decision-tree, is capable to classify the histotype of the hepatic lesions, based on their CT quantitative and qualitative features, with a moderate overall accuracy (62%). The accuracy in the classification of the histotypes ranged between 0.53 for nodular hyperplasia, and 0.92 for the benign lesions other than the nodular hyperplasia. However, the dataset wasn't divided into training-set and test-set due to its poor dimensions. Therefore, the same dataset used for training the decision-tree has been used for the calculation of its accuracy.

Furthermore, we developed a decision-tree for the classification of malignant focal liver based on their CEUS features. The model successfully classified the new cases with an accuracy of 0.79. All the sarcoma cases have properly detected and the hepatocarcinoma cases have been accurately identify. By the way, in this case benign lesions were not taken into account. The decision-tree is an easy and useful tool in everyday clinical practice for helping the veterinarian in the formulation of a diagnostic suspect.

In our metanalysis study, all the descriptive studies focused on CT features of focal liver lesions have been reviewed. Some features resulted to be useful in the distinction between benign and malignant lesions. Specifically, the presence of a capsule, hypoattenuation and heterogeneity of the lesion in delayed phase showed high DOR values, therefore could be used for focal liver lesions classification. In addition, other features such as maximum dimension, attenuation of the radiologically normal parenchyma during pre-contrast, arterial, portal and delayed phase resulted to be significantly different among benign and malignant lesions. This an interesting result, suggesting that also radiologically normal liver parenchyma should be sampled during cytological or histological examination.

For the focal splenic lesions, in our study we have substantially confirmed what already reported in the literature. Indeed, the results of the study suggested that sarcomas are larger compared to nodular hyperplasia, and by a cyst-like appearance. The developed decision-tree showed high accuracy in the classification of sarcomas but low for benign lesions. More specifically, a cut-off of 3.6 cm was identified to distinguish sarcoma from nodular hyperplasia. Similarly, a cut-off value of 44 HU was for the distinction between benign lesions other than nodular hyperplasia and sarcomas wa identified. Furthermore, a 126 HU cut-off for the distinction between nodular hyperplasia (< 126 HU) and other benign lesions (> 126 HU) was identified. The decision-tree can be considered a viable and simple tool for the veterinary practise in this case too.

In the second phase of our study, we have trained and tested a CNN (Inception V3) on the CT images of focal hepatic and splenic lesions. The applications of CNNs to diagnostic imaging is giving promising results in human medicine where the CNNs are trained on data-set considerably bigger (hundred of thousand of cases) compared to the data-set available in our study (69 cases for liver and 52 for spleen).

Several methods of data augmentation can partially overcome the limitations deriving from small datasets in the reduction of overfitting. In this project we applied cropping, translation, flipping, rotation and contrast change methods of data augmentation. Indeed, these are highlighted as the most effective in medical imaging assessment [9].

Despite the limited results of the study in the application of CNNs on hepatic and splenic pathologies obtained in this study, I am keen on improving the training the CNN on focal liver and splenic lesions increasing as much as possible the record of cases.

Hepatic and splenic pathologies are a complex matter and the there is an extremely wide range of a benign or malignant possible differential diagnosis and to have a concrete and reliable diagnosis in short terms is a focal target in veterinary medicine leading to promptly approach the case with the best possible therapy. CT is one of the main diagnostic imaging techniques used veterinary medicine and, with our research, the results of this PhD have provided some tools that could be employed in the everyday veterinary clinical practice.

BIBLIOGRAPHY

1. Stehlík L, Di Tommaso M, Signore F Del, et al. Triple-phase multidetector computed tomography in distinguishing canine hepatic lesions. *Animals*. 2021;11(1):1-12. doi:10.3390/ani11010011
2. Burti S, Zotti A, Bonsembiante F, Contiero B, Banzato T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. *Front Vet Sci*. 2021;8(March):1-10. doi:10.3389/fvets.2021.611556
3. Leela Arporn R, Ohta H, Shimbo G, et al. Computed tomographic features for differentiating benign from malignant liver lesions in dogs. *J Vet Med Sci*. 2019;81(12):1697-1704. doi:10.1292/jvms.19-0278
4. Griebie ER, David FH, Ober CP, et al. Evaluation of canine hepatic masses by use of triphasic computed tomography and B-mode, color flow, power, and pulsed-wave Doppler ultrasonography and correlation with histopathologic classification. *Am J Vet Res*. 2017;78(11):1273-1283.
5. Jones ID, Lamb CR, Drees R, Priestnall SL, Mantis P. Associations between dual-phase computed tomography features and histopathologic diagnoses in 52 dogs with hepatic or splenic masses. *Vet Radiol Ultrasound*. 2016;57(2):144-153. doi:10.1111/vru.12336
6. Kutara K, Seki M, Ishikawa C, et al. Triple-phase helical computed tomography in dogs with hepatic masses. *Vet Radiol Ultrasound*. 2014;55(1):7-15. doi:10.1111/vru.12099
7. Fukushima K, Kanemoto H, Ohno K, et al. Ct characteristics of primary hepatic mass lesions in dogs. *Vet Radiol Ultrasound*. 2012;53(3):252-257. doi:10.1111/j.1740-8261.2011.01917.x
8. Taniura T, Marukawa K, Yamada K, Hikasa Y, Ito K. Differential diagnosis of hepatic tumor-like lesions in dog by using dynamic CT scanning. *Hiroshima J Med Sci*. 2009;58(1):17-24.
9. Sahiner B, Pezeshk A, Hadjiiski LM, et al. Deep learning in medical imaging and radiation therapy. *Med Phys*. 2019;46(1):e1-e36. doi:10.1002/mp.13264