

Detecting adrenal lesions using machine learning - a state of the art

Bernardo Gonçalves - 58885 - Doctoral program in Biomedical Engineering

December 17, 2022

1 Adrenal Glands and Adrenal Lesions

1.1 Anatomy, physiology and pathophysiology (CHANGE)

The adrenal glands, or suprarenal glands, are two small glands located on top of the kidneys. Each has a body and two limbs [1], and weights about 5g [2]. Figure 1 shows the localisation and anatomy of the adrenal glands. A slice of the right adrenal gland can be linear, comma or V-shaped. On the other hand, the left adrenal is triangular or Y-shaped.

The adrenal gland has two different regions: the adrenal cortex and the adrenal medulla. The adrenal medulla is the inner region of the gland and is derived from the mesoderm [1]. This region is responsible for the production of epinephrine and norepinephrine, which stimulates the fight-or-flight response [?]. The adrenal cortex is the outer region of the gland and is derived from neural crest cells [1]. This region is responsible for the production of cortisol, corticosterone, and cortisone, which increase blood glucose levels, and the production of aldosterone, which increases the level of sodium in the blood [?].

The adrenal glands can be affected by a wide variety of benign and malignant lesions. It is estimated that approximately 6% of the population has them [3]. These lesions can be primary if they originated in the glands themselves (cortex or medulla) or secondary if they have another origin. Primary lesions can be functional if they produce hormones [3]. Table 1 presents an overview of the adrenal lesions. The most common adrenal lesions are adenomas. Adenomas are often non-functional and remain asymptomatic, being discovered incidentally [?]. Adenomas are, in most cases, non-functional and remain asymptomatic, being discovered incidentally [4]. Adrenocortical carcinoma is a rare lesion although it is the most common primary malignant adrenal lesion. This lesion affects children in their first decade and adults in their fourth and fifth decades [3]. Also, the adrenals are a frequent location of metastases [4].

Functional lesions can cause endocrines syndromes, such as Conn and Cushing syndrome. The Cushing syndrome or hypercortisolism is caused by elevated values of cortisol and it is associated with adrenal adenomas, mostly. Nevertheless, adrenocortical carcinomas or pheochromocytomas can also cause Cushing syndrome. This syndrome is defined by symptoms like obesity, rounded face, abnormal skin pigmentation, muscle weakness, hypertension, diabetes, and others. On the other hand, the Conn syndrome or primary aldosteronism is related to the excessive production of aldosterone. The most common symptoms of this syndrome are sodium retention, plasma renin suppression, hypertension, cardiovascular damage, and increased potassium excretion. Like the Conn syndrome, this syndrome is commonly caused by adrenal adenomas. In opposition to the above-mentioned syndromes, the Addison disease is caused by adrenal insufficiency, which can be caused by malignant lesions. Patients with this disease can experience weight loss, weakness, fatigue, gastrointestinal upset, orthostatic hypotension, and abnormal skin pigmentation. These symptoms can evolve into dehydration, shock, hyperkalaemia (high potassium), and hyponatremia (high sodium) when entering acute adrenal insufficiency [?].

1.2 Adrenals Imaging

Structural medical imaging techniques are decisive to detect and characterize adrenal lesions and complementary to functional imaging and endocrine evaluation in the assessment of functional lesions. Imaging techniques can also rule out invasive interventions. The most used imaging techniques to evaluate the adrenal glands are Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). The Ultrasonography (USG) despite being a common method to assess abdominal pathologies, is not a good method to perceive retroperitoneal (back of the peritoneum) structures

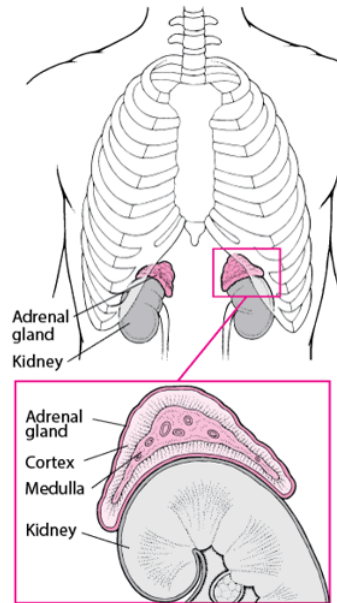
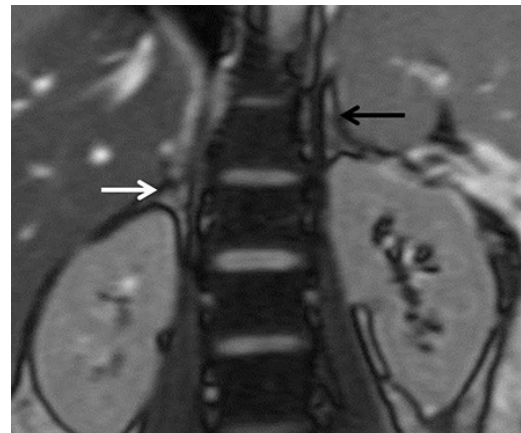


Figure 1: Adrenal localization and anatomy. From **ADD REFERENCE**



(a) Adrenal glands in a contrast-enhanced CT axial slice in arterial phase. Due to the high level of retroperitoneal fat both glands are enhanced in this image slice. Reprinted from [3]



(b) Adrenal glands in a MR CSI coronal slice. Both glands have a intermediate signal intensity. Reprinted from [3]

Figure 2: Both figures show the y and v shaped adrenal glands. The arrows indicate the localization of the glands.

like the adrenals [3]. Both Figures [ref FIG 2] and [REF FIG 3] show the V and Y-shaped normal glands. Figure [ref FIG 2] is an axial contrast-enhanced CT image in the arterial phase where both glands are enhanced due to the high retroperitoneal fat content. Figure [REF FIG 3] is a coronal MR Chemical Shift Image (CSI) out-of-phase showing normal adrenal glands also.

Commonly, non-functional lesions do not require any treatment and, for that reason it is crucial to differentiate between adenomas and non-adenomas [4]. Adrenal adenomas have less than 1 cm in diameter, usually and they can be lipid-rich or lipid-poor [3]. About 70-80 % of the adenomas are lipid-rich in contrast with the malignant lesions [4]. This results in a 20-30 % overlap between adenomas and malignant lesions in terms of intracytoplasmic lipid content [5]. Lipid-rich adenomas can be easily detected using unenhanced CT (less than 10 HU) [3] or CSI [4]. However, lipid-poor adenomas cannot be correctly characterised by unenhanced CT [5]. In these cases, CSI presents itself as a better solution because of its improved sensitivity to low levels of lipid content and therefore it can detect 62-67% of the adenomas uncharacterised by CT [5]. CSI is a fat-suppression technique that originates two sets of images: in-phase (IP) and out-of-phase (OP) images. In OP images the signal is the difference between the signals of water and fat molecules. In IP images the signal of both water and fat is added. Thus, there is a significant suppression of signal from IP to OP images in lipid-rich lesions [6]. OP images are characterised by the so-called India ink artefact that is a signal void in the margins of fatty and normal tissues [6], creating a darker boundary in lipid-rich lesions like most adenomas. [4] performed a metanalysis with 1280 lesions (859 adenomas, 421 non-adenomas) and documented a sensitivity of 94% and specificity of 95% detecting adrenal adenomas. The same study states the importance of using Dynamic Contrast-Enhanced (DCE) MRI to improve the detection performance of lipid-poor adenomas. Adenomas present a signal increase in the arterial phase and a rapid washout [7].

2 Differential diagnosis of adrenal lesions - current medical approach

2.1 Current procedure

2.2 Main limitations and setbacks

3 Differential diagnosis of adrenal lesions - machine learning approach

3.1 Research details

The studies analysed in this paper were selected accordingly in the following PICO criteria:

P (patients) – patients with adrenal lesions.

I (interventions) – machine learning (including deep learning) modelling.

C (comparison) – standard of care imaging including Computed Tomography (CT) and Magnetic Imaging Resonance (MRI).

O (outcome) – lesion differentiation (benign/malign and subtyping) and lesions detection.

The studies were searched in two databases: PubMed and Web of Science in September 2022. The search was made with the following research string: (adrenal or suprarenal) AND (CT OR "computed tomography" OR MRI OR "magnetic resonance imaging" OR "MRI scan" OR "nuclear magnetic resonance" OR "magnetic resonance" OR NMR) AND ("deep learning" OR "convolutional networks" OR CNN OR "neural networks" OR convolutional OR DNN OR SVM OR "Support vector machine" OR "decision tree" OR "machine learning"). Studies that were: (a) reviews, (b) not written in English, (c) did not report a modelling method, (d) did not have the full text available were excluded from this research.

The research resulted in 23 studies that were divided in 3 groups related with their object of study:

Group A: contains all the studies that focus on the differentiation between adrenal adenomas and other adrenal lesions.

Group B: contains all the studies that target the differentiation between benign and malignant adrenal lesions.

Group	Image Modality			Model Type			Total
	MRI	CT	MRI+CT	ML	DL	ML+DL	
A - Adenomas vs other lesions	4	7	1	11	1	0	12
B - Benign vs Malign	2	5	0	7	0	0	7
C - Other	1	3	0	2	1	1	4
Sum	7	15	1	20	2	1	23

Table 1: Studies distribution per group. CT: Computed Tomography; MRI: Magnetic Resonance Imaging; ML: Traditional Machine Learning; DL: Deep Learning.

Group C: contains the remaining studies that did not fit any of the above categories.

Table 1 shows the distribution of the selected studies in terms of their group, image modality and model type. Overall, most of the studies adopt traditional machine learning models to classify imaging features (radiomics) from CT images to distinguish adenomas from other types of adrenal lesions, eg. Metastases, pheochromocytomas.

Radiomics is the extraction of quantitative features from medical images. Before the feature extraction, there are two relevant steps: image acquisition and segmentation of the region of interest (ROI). Any type of medical images can be used in radiomics, such as Positron Emission Tomography (PET), MRI, or CT. The ROI segmentation is necessary to limit the amount of data that needs to be processed in order to extract features. This process can be completely manual, which is the gold standard, where a specialist selects the ROI. Manual segmentation is a time-consuming task that significantly depends on the skill of the operator. There are fully automatic methods for ROI segmentation however, they can fail in difficult cases, such as, lesions with indistinct borders and are highly dependent on the quality of the image. For that reason, the usage of semi-automatic methods is preferable. These methods have minimal user interaction (seed identification or manual correction). The extracted quantitative features aim to describe the complexity of the individual region of interest. Ordinarily, these features are divided into 4 categories:

1. Shape-Based Features: numeric information respecting geometrics characteristics, like shape and size.
2. First-Order Statistics: distribution of voxel values without spatial information, generally histogram-based.
3. Second-Order Statistics: “texture” features, focus on the spatial relationships between voxels with similar grey levels.
4. High-Order Statistics: usage of filters to extract patterns from the images. From the resultant images, first and second-order features are extracted.

The most relevant radiomic features to the task in hand are selected using statistical approaches or machine learning [8]. Then, these features are used as the input of ML models to classify the region of interest. This type of workflow is widespread, appearing in 87% (20/23) of the analysed studies. Traditional machine learning models like K-means, Support Vector Machine (SVM), Logistic Regression (LogReg), or Random Forest (RanFor), are frequently used to classify radiomic features of regions of interest and have achieved high performance in different anatomical regions [8, 9]. Deep Learning models, such as, Convolutional Neural Networks (CNNs), have been often applied to medical images from several anatomical regions with promising results [10]. Despite that, only 13% of the studies presented in this paper report the application of DL models to adrenal images. DL models are different from ML models mostly because they demand bigger annotated data sets and they do not rely on the feature extraction step (all features are automatically extracted and classified by the model).

In the next sections each group of papers will be analysed in detail, exploring their common and contrasting aspects. For each group the analysis will be focused on the utilized data sets, models, and the obtained results.

3.2 Group A - Adenomas vs other lesions

Group A comprises all studies that focus on the differentiation between adenomas and other lesions. This group corresponds to 52% of all analysed studies. Tables 2, 3 and 4 present an overview

Reference	Image Modality	Sample Size (lesions)		
		Total	Adenomas	Other
[12]	U-CT	76	36	40
[14]	U/CE-CT	110	80	30
[15]	U-CT	265	181	84
[18]	CE-CT	54	25	29
[19]	U/CE-CT	19	9	10
[22]	U/CE-CT	115	83	32
[16]	U/CE-CT	280	188	92
[20]	U/CE-CT; T1W-OP/IP MRI	23	15	8
[17]	T1W-OP/IP; T2W MRI	60	40	20
[11]	T1W-OP/IP; T2W MRI	44	29	15
[13]	T1W-OP/IP; T2W MRI	63	23	40
[21]	T1W-OP/IP; T2W MRI	60	40	20

Table 2: Dataset Details for each article in the Group A. CT: Computed Tomography; U: Un-enhanced; CE: Contrast Enhanced; MRI: Magnetic Resonance Imaging; OP: Out-of-phase; IP: In-phase; T1W: T1-weighted; T2W: T2-weighted.

of the data sets, models, and results, respectively, for each study within group A. [11, 12, 13] have data sets with adenomas and metastases. [14, 15, 16, 17] have patients with adenomas and pheochromocytomas. [18, 19, 20] have compared adenomas and carcinomas. [13, 15, 14, 16] have considered only lipid-poor adenomas for their adenoma dataset. Finally, [21] data set has 3 classes: lipid-poor adenomas, lipid-rich adenomas and nonadenomas and [22] performs a general differentiation between adenomas and nonadenomas. In terms of the imaging modality utilised by the studies, there is slight advantage to CT, as it is used in 8 studies, mostly with and without contrast enhanced images. The studies that analyse MRI images focus on chemical shift images as they pose an advantage in the detection of the lipid content of the lesions. The sample size of the studies ranges from dozens to hundreds of lesions, which is closely related to the applied inclusion criteria and the initial sample size. For example, [19] had a database of 336 patients however only 19 met the inclusion criteria, each with one lesion.

All studies, but one, performed lesion classification with ML models using radiomic features. The most common model is logistic regression (LogReg), followed by support vector machine (SVM) and decision tree-based models. In every study, the region of interest (ROI) was manually selected by experts. The extraction of first and second-order statistics is a common practice, but only 3 studies extracted shape-based features, and none extracted higher-order statistics. The only study that has implemented a DL model has performed ROI (selected, cropped and labelled by experts) classification with a deep convolution neural network [22]. They report the usage of augmentation techniques such as rotations and horizontal flips. [19] was the only study to implement an unsupervised model.

In [20] the data set consists of lipid-poor adenomas and carcinomas in both MRI and CT images. The objective of this work was to compare the different image modalities using the same machine learning approach. The authors have reported only the Area Under the Receiving Operating Characteristic Curve (AUC). The value presented in 4 refers to the best result, using CE-CT images, with MRI images the value decreases to 58 %. Both studies by Yi et al [15, 14] implemented a logistic regression model with the same radiomic features but using different CT images and achieved impressive results. [15] adds clinical features such as, the existence of necrosis or calcification and lesion dimensions, to the radiomic features. Also, to improve feature selection the authors used the Least Absolute Shrinkage and Selection Operator (LASSO).

3.3 Group B - Malign vs benign lesions

Group B consists of studies that aim at the differentiation between benign and malignant lesions. This group includes 30% of the studies. Tables 5, 6 and 7 display an overview of the data sets, models, and results, respectively, for each study within group B. In terms of the image modalities in the data sets, this group is similar to group A. There are more studies that use CT images and all of them, except one, use CE-CT. Studies that analyse MRI data sets have both T1W chemical shift images and T2W images. [23] has the largest data set of all the analysed studies. Group B does not comprise any study with a DL model, however there are two studies that use neural

Reference	Type	Classification Model	ROI	Features
[12]	ML	LogReg	Manual	1 st
[14]	ML	LogReg	Manual	1 st , 2 nd , higher
[15]	ML	LogReg	Manual	1 st , 2 nd , higher
[18]	ML	RanFor; LogReg	Manual	1 st , 2 nd , shape
[19]	ML	K-Means	Manual	1 st , 2 nd
[22]	DL	DCNN	Manual	-
[16]	ML	LinReg; SVM; RanFor	Manual	1 st ; clinical
[20]	ML	LogReg	Manual	1 st , 2 nd , shape
[17]	ML	SVM	Manual	1 st
[11]	ML	LogReg	Manual	1 st
[13]	ML	LogReg	Manual	1 st , shape
[21]	ML	DecTre	Manual	1 st , 2 nd

Table 3: Modelling Details for each article in the Group A. ML: Traditional Machine Learning models; DL: Deep Learning models; LogReg: Logistic Regression; LASSO: Least Absolute Shrinkage and Selection Operator; DecTre: Decision Tree; RanFor: Random Forest; PCA: Principal Components Analysis; SVM: Support Vector Machine; 1st, 2nd, higher: first, second, higher order statistics, respectively; shape: shape-based features.

Reference	Specificity - %	Sensitivity - %	Accuracy - %	AUC - %
[12]	75.0	47.5	60.5	65.0
[14]	97.5	86.2	94.4	95.2
[15]	90.3	95.5	92.0	95.7
[18]	83.0	81.0	82.0	89.0
[19]	90.0	87.5	88.9	-
[22]	96.0	87.0	94.0	-
[16]	86.6	89.2	87.5	-
[20]	-	-	80.0	-
[17]	-	-	85.0	91.7
[11]	86.2	93.3	88.6	97.0
[13]	100	75.0	84.1	-
[21]	-	-	80.0	-

Table 4: Model metrics for each article in the Group A. AUC: Area Under the Receiving Operating Characteristic Curve.

Reference	Image Modality	Sample Size (lesions)		
		Total	Benign	Malign
[23]	U-CT	377	182	195
[24]	CE-CT	114	90	24
[26]	U/CE-CT	210	114	96
[27]	CE-CT	160	89	71
[28]	U/CE-CT	40	21	19
[25]	T1W-OP/IP; T2W MRI	114	9	105
[29]	T1W-OP/IP; T2W MRI	55	37	18

Table 5: Dataset Details for each article in the Group B. CT: Computed Tomography; U: Un-enhanced; CE: Contrast Enhanced; MRI: Magnetic Resonance Imaging; OP: Out-of-phase; IP: In-phase; T1W: T1-weighted; T2W: T2-weighted.

Reference	Type	Classification Model	ROI	Features
[23]	ML	LogReg	-	1 st , 2 nd
[24]	ML	NN	Semi-auto	1 st , 2 nd , higher, shape
[26]	ML	BayCla	Semi-auto	2 nd
[27]	ML	LogReg	Semi-auto	1 st , higher
[28]	ML	RanFor	Manual	1 st , 2 nd , higher, shape
[25]	ML	SVM	Manual; Semi-auto	2 nd , higher
[29]	ML	ExTre	Manual	1 st , 2 nd , higher, shape

Table 6: Modelling Details for each article in the Group B. ML: Traditional Machine Learning models; DL: Deep Learning models; LogReg: Logistic Regression; BayCla: Bayesian Classifier; NN: Neural Network; ExTre: Extra Trees Classifier. RanFor: Random Forest; SVM: Support Vector Machine; 1st, 2nd, higher: first, second, higher order statistics, respectively; shape: shape-based features.

networks. [24] experimented several optimisation algorithms for neural networks and achieved the best results using the Bounded Particle Swarm Optimisation algorithm ¹. In [25] a SVM was implemented to perform binary classification, however, the authors also used a NN to perform type characterisation. The authors divided the dataset in 4 classes, each with one type of lesion, 3 benign (adenoma, cyst and lipoma) and 1 malign (metastasis). For both workflows, ROI selection was made using manual and semi-automatic segmentation, and the same radiomic features were used. The results presented in Table 7 refer to the binary classification, and they were the best results of this group. For the multiclass classification the results were poor, despite the high values of specificity and accuracy, 96.2 % and 93.2 %, respectively, the sensibility is extremely low, 59.6 %, which can be explained by the high number of class and the lack of balance in the data set. In this group all the models are supervised learning models.

¹https://link.springer.com/chapter/10.1007/978-3-319-93025-1_2

Reference	Specificity - %	Sensitivity - %	Accuracy - %	AUC - %
[23]	-	-	-	78.0
[24]	82.2	75.0	80.7	78.6
[26]	67.5	94.8	80.0	-
[27]	77.0	58.0	68.0	73.0
[28]	71.4	84.2	77.5	85.1
[25]	90.0	99.2	98.4	-
[29]	-	-	91.0	97.0

Table 7: Model metrics for each article in the Group B. AUC: Area Under the Receiving Operating Characteristic Curve.

Reference	Image Modality	Sample size	Task
[30]	U-CT	38	Lesion Detection
[31]	CE-CT	229	Multiclass Classification
[32]	T2W-MRI	305	pheo vs non-pheo
[33]	U/CE-CT	83	Adenoma subtyping

Table 8: Dataset Details for each article in the Group C. CT: Computed Tomography; U: Un-enhanced; CE: Contrast Enhanced; MRI: Magnetic Resonance Imaging; OP: Out-of-phase; IP: In-phase; T1W: T1-weighted; T2W: T2-weighted. Pheo: pheochromocytomas

Reference	Type	Model	ROI	Features
[30]	DL	FCN	Manual	-
[31]	DL + ML	CNN + SVM	Manual	CNN embedding
[32]	ML	LogReg	Semi-Auto	1 st , 2 nd , higher, shape
[33]	ML	LogReg	Manual	1 st , higher, shape

Table 9: Modelling Details for each article in the Group C. ML: Traditional Machine Learning models; DL: Deep Learning models; LogReg: Logistic Regression; BSGC: Bayesian Spatial Gaussian Classifiers; CNN: Convolutional Neural Network; SVM: Support Vector Machine; 1st, 2nd, higher: first, second, higher order statistics, respectively; shape: shape-based features.

3.4 Group C

Group C consists of the remaining studies that did not fit any of prior defined groups. This group includes 18% of the studies. Tables 5, 6 and 7 display an overview of the data sets, models, and results, respectively, for each study inside group C. In this group are 4 studies with distinct objectives. [30] implements a fully convolutional neural network for lesion detection using a small dataset of U-CT images. [31] creates a machine learning pipeline where the CNN embedding is used as input of an SVM to execute multiclass classification with a data set of CE-CT images. The data set had 5 classes: carcinoma, non-functional adenoma, ganglioneuroma, myelolipoma and pheochromocytoma.[32] aimed at the differentiation between pheochromocytomas and non-pheochromocytomas with a T2W MRI dataset and a logistic regression model. [33] also used logistic regression but to characterize adenomas with a CT dataset.

References

- [1] Y. Baba and O. Bashir, “Adrenal gland,” *Radiopaedia.org*, 1 2012.
- [2] “adrenal gland — definition, anatomy, and function — britannica,” 2012.
- [3] A. Panda, C. J. Das, E. Dhamija, R. Kumar, and A. K. Gupta, “Adrenal imaging (part 1): Imaging techniques and primary cortical lesions,” *Indian Journal of Endocrinology and Metabolism*, vol. 19, pp. 8–15, 1 2015.

Reference	Specificity - %	Sensitivity - %	Accuracy - %	AUC - %
[30]	-	76.29	-	-
[31]	95.9	83.7	85.2	-
[32]	75.0	85.7	84.0	90.6
[33]	92.8	91.5	92.2	90.2

Table 10: Model metrics for each article in the Group C. AUC: Area Under the Receiving Operating Characteristic Curve.

- [4] I. Platzek, D. Sieron, V. Plodeck, A. Borkowetz, M. Laniado, and R. T. Hoffmann, “Chemical shift imaging for evaluation of adrenal masses: a systematic review and meta-analysis,” 2019.
- [5] G. M. Israel, M. Korobkin, C. Wang, E. N. Hecht, and G. A. Krinsky, “Comparison of unenhanced ct and chemical shift mri in evaluating lipid-rich adrenal adenomas,” *American Journal of Roentgenology*, vol. 183, pp. 215–219, 2004.
- [6] V. Jahanvi and A. Kelkar, “Chemical shift imaging: An indispensable tool in diagnosing musculoskeletal pathologies,” *SA Journal of Radiology*, 2021.
- [7] J. J. Chung, R. C. Semelka, and D. R. Martin, “Adrenal adenomas: characteristic post-gadolinium capillary blush on dynamic mr imaging.,” *Journal of magnetic resonance imaging : JMRI*, vol. 13, pp. 242–8, 2 2001.
- [8] H. Zhang, H. Lei, and J. Pang, “Diagnostic performance of radiomics in adrenal masses: A systematic review and meta-analysis,” *Frontiers in Oncology*, vol. 12, 9 2022.
- [9] M. W. Wagner, K. Namdar, A. Biswas, S. Monah, F. Khalvati, and B. B. Ertl-Wagner, “Radiomics, machine learning, and artificial intelligence—what the neuroradiologist needs to know,” *Neuroradiology*, vol. 63, pp. 1957–1967, 12 2021.
- [10] A. Anaya-Isaza, L. Mera-Jiménez, and M. Zequera-Diaz, “An overview of deep learning in medical imaging,” *Informatics in Medicine Unlocked*, vol. 26, p. 100723, 1 2021.
- [11] N. Schieda, S. Krishna, M. D. McInnes, B. Moosavi, A. Alrashed, R. Moreland, and E. S. Siegelman, “Utility of mri to differentiate clear cell renal cell carcinoma adrenal metastases from adrenal adenomas,” *AJR. American journal of roentgenology*, vol. 209, pp. W152–W159, 9 2017.
- [12] W. Tu, R. Verma, S. Krishna, M. D. McInnes, T. A. Flood, and N. Schieda, “Can adrenal adenomas be differentiated from adrenal metastases at single-phase contrast-enhanced ct?,” <https://doi.org/10.2214/AJR.17.19276>, vol. 211, pp. 1044–1050, 9 2018.
- [13] W. Tu, J. Abreu-Gomez, A. Udare, A. Alrashed, and N. Schieda, “Utility of t2-weighted mri to differentiate adrenal metastases from lipid-poor adrenal adenomas,” *Radiology: Imaging Cancer*, vol. 2, 11 2020.
- [14] X. Yi, X. Guan, C. Chen, Y. Zhang, Z. Zhang, M. Li, P. Liu, A. Yu, X. Long, L. Liu, B. T. Chen, and C. Zee, “Adrenal incidentaloma: machine learning-based quantitative texture analysis of unenhanced ct can effectively differentiate spheo from lipid-poor adrenal adenoma,” *Journal of Cancer*, vol. 9, pp. 3577–3582, 2018.
- [15] X. Yi, X. Guan, Y. Zhang, L. Liu, X. Long, H. Yin, Z. Wang, X. Li, W. Liao, B. T. Chen, and C. Zee, “Radiomics improves efficiency for differentiating subclinical pheochromocytoma from lipid-poor adenoma: a predictive, preventive and personalized medical approach in adrenal incidentalomas,” *EPMA Journal*, vol. 9, pp. 421–429, 12 2018.
- [16] H. Liu, X. Guan, B. Xu, F. Zeng, C. Chen, H. L. Yin, X. Yi, Y. Peng, and B. T. Chen, “Computed tomography-based machine learning differentiates adrenal pheochromocytoma from lipid-poor adenoma,” *Frontiers in Endocrinology*, vol. 13, 3 2022.
- [17] J. Liu, K. Xue, S. Li, Y. Zhang, and J. Cheng, “Combined diagnosis of whole-lesion histogram analysis of t1- and t2-weighted imaging for differentiating adrenal adenoma and pheochromocytoma: A support vector machine-based study,” *Canadian Association of Radiologists Journal*, vol. 72, pp. 452–459, 8 2021.
- [18] M. M. Elmohr, D. Fuentes, M. A. Habra, P. R. Bhosale, A. A. Qayyum, E. Gates, A. I. Morshid, J. D. Hazle, K. M. Elsayes, A. M. I, J. D. Hazle, and K. M. Elsayes, “Machine learning-based texture analysis for differentiation of large adrenal cortical tumours on ct,” *CLINICAL RADIOLOGY*, vol. 74, pp. 818.e1–818.e7, 10 2019.
- [19] F. Torresan, F. Crimi, F. Ceccato, F. Zavan, M. Barbot, C. Lacognata, R. Motta, C. Armellin, C. Scaroni, E. Quaia, C. Campi, and M. Iacobone, “Radiomics: a new tool to differentiate adrenocortical adenoma from carcinoma,” *BJS open*, vol. 5, 1 2021.

- [20] L. M. Ho, E. Samei, M. A. Mazurowski, Y. Zheng, B. C. Allen, R. C. Nelson, and D. Marin, "Can texture analysis be used to distinguish benign from malignant adrenal nodules on unenhanced ct, contrast-enhanced ct, or in-phase and opposed-phase mri?," *American Journal of Roentgenology*, vol. 212, pp. 554–561, 3 2019.
- [21] V. Romeo, S. Maurea, R. Cuocolo, M. Petretta, P. P. Mainenti, F. Verde, M. Coppola, S. Dell'Aversana, and A. Brunetti, "Characterization of adrenal lesions on unenhanced mri using texture analysis: A machine-learning approach," *Journal of Magnetic Resonance Imaging*, vol. 48, pp. 198–204, 7 2018.
- [22] M. Kusunoki, T. Nakayama, A. Nishie, Y. Yamashita, K. Kikuchi, M. Eto, Y. Oda, and K. Ishigami, "A deep learning-based approach for the diagnosis of adrenal adenoma: a new trial using ct," *The British Journal of Radiology*, vol. 95, 7 2022.
- [23] K. Shoemaker, B. P. Hobbs, K. Bharath, C. S. Ng, and V. Baladandayuthapani, "Tree-based methods for characterizing tumor density heterogeneity," *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, vol. 23, pp. 216–227, 1 2018.
- [24] H. Koyuncu, R. Ceylan, S. Asoglu, H. Cebeci, and M. Koplay, "An extensive study for binary characterisation of adrenal tumours.," *Medical & Biological Engineering & Computing*, vol. 57, pp. 849–862, 4 2019.
- [25] M. Barstugan, R. Ceylan, S. S. S. Asoglu, H. Cebeci, and M. Koplay, "Adrenal tumor characterization on magnetic resonance images," *International Journal of Imaging Systems and Technology*, vol. 30, pp. 252–265, 3 2020.
- [26] X. Li, M. Guindani, C. S. Ng, and B. P. Hobbs, "Spatial bayesian modeling of glm with application to malignant lesion characterization," *Journal of applied statistics*, vol. 46, pp. 230–246, 1 2019.
- [27] M. B. Andersen, U. Bødtger, I. R. Andersen, K. S. Thorup, B. Ganeshan, and F. Rasmussen, "Metastases or benign adrenal lesions in patients with histopathological verification of lung cancer: Can ct texture analysis distinguish?," *European Journal of Radiology*, vol. 138, p. 109664, 5 2021.
- [28] A. W. Moawad, A. Ahmed, D. T. Fuentes, J. D. Hazle, M. A. Habra, and K. M. Elsayes, "Machine learning-based texture analysis for differentiation of radiologically indeterminate small adrenal tumors on adrenal protocol ct scans," *Abdominal Radiology*, vol. 46, pp. 4853–4863, 10 2021.
- [29] A. Stanzione, R. Cuocolo, F. Verde, R. Galatola, V. Romeo, P. P. Mainenti, G. Aprea, E. Guadagno, M. D. B. D. Caro, and S. Maurea, "Handcrafted mri radiomics and machine learning: Classification of indeterminate solid adrenal lesions," *Magnetic Resonance Imaging*, vol. 79, pp. 52–58, 6 2021.
- [30] L. Bi, J. Kim, T. Su, M. Fulham, D. Feng, and G. Ning, "Adrenal lesions detection on low-contrast ct images using fully convolutional networks with multi-scale integration," pp. 895–898, IEEE Computer Society, 6 2017.
- [31] L. Bi, J. Kim, T. Su, M. Fulham, D. D. Feng, and G. Ning, "Deep multi-scale resemblance network for the sub-class differentiation of adrenal masses on computed tomography images," *Artificial Intelligence in Medicine*, vol. 132, p. 102374, 10 2022.
- [32] J. Kong, J. Zheng, J. Wu, S. Wu, J. Cai, X. Diao, W. Xie, X. Chen, H. Yu, L. Huang, H. Fang, X. Fan, H. Qin, Y. Li, Z. Wu, J. Huang, and T. Lin, "Development of a radiomics model to diagnose pheochromocytoma preoperatively: a multicenter study with prospective validation," *Journal of Translational Medicine*, vol. 20, 12 2022.
- [33] Y. Zheng, X. Liu, Y. Zhong, F. Lv, and H. Yang, "A preliminary study for distinguish hormone-secreting functional adrenocortical adenoma subtypes using multiparametric ct radiomics-based machine learning model and nomogram," *Frontiers in Oncology*, vol. 10, 9 2020.