

Analyzing Tumors by Synthesis

Qi Chen, Yuxiang Lai, Xiaoxi Chen, Qixin Hu, Alan Yuille and Zongwei Zhou

Abstract Computer-aided tumor detection has shown great potential in enhancing the interpretation of over 80 million CT scans performed annually in the United States. However, challenges arise due to the rarity of CT scans with tumors, especially early-stage tumors. Developing AI with real tumor data faces issues of scarcity, annotation difficulty, and low prevalence. Tumor synthesis addresses these challenges by generating numerous tumor examples in medical images, aiding AI training for tumor detection and segmentation. Successful synthesis requires realistic and generalizable synthetic tumors across various organs. This chapter reviews AI development on real and synthetic data and summarizes two key trends in synthetic data for cancer imaging research: modeling-based and learning-based approaches. Modeling-based methods, like Pixel2Cancer, simulate tumor development over time using generic rules, while learning-based methods, like DiffTumor, learn from a

Qi Chen
University of Chinese Academy of Sciences
e-mail: chenqi24@ucas.ac.cn

Yuxiang Lai
Emory University
e-mail: ylai76@emory.edu

Xiaoxi Chen
University of Illinois Urbana-Champaign
e-mail: xiaoxic3@illinois.edu

Qixin Hu
University of Southern California
e-mail: qixinhu@usc.edu

Alan Yuille
Johns Hopkins University
e-mail: ayuille1@jh.edu

Zongwei Zhou (✉)
Johns Hopkins University
e-mail: zzhou82@jh.edu

few annotated examples in one organ to generate synthetic tumors in others. Reader studies with expert radiologists show that synthetic tumors can be convincingly realistic. We also present case studies in the liver, pancreas, and kidneys reveal that AI trained on synthetic tumors can achieve performance comparable to, or better than, AI only trained on real data. Tumor synthesis holds significant promise for expanding datasets, enhancing AI reliability, improving tumor detection performance, and preserving patient privacy.

1 Introduction

Medical image analysis aims to derive detailed information non-invasively about a patient’s medical condition, including the disease’s origin, precise location, and its relationship with adjacent tissues. Specifically, for symptoms that cannot be directly diagnosed, medical professionals employ various imaging devices to capture detailed images of target organs for disease screening, diagnosis, and treatment. Therefore, medical imaging systems generate vast amounts of medical image data daily. This data may encompass different organs of the body, as well as tissues and pathological regions associated with diseases.

Medical images come from modalities like X-ray, CT, MRI, and PET. This data, referred to as **real data** in this chapter, can be analyzed using post-processing and artificial intelligence (AI) to reveal details not visible to the naked eye, aiding in disease detection such as detecting tumors at their early stage. However, managing real-world data for AI-driven diagnostics is challenging. **Synthetic data** offers a promising alternative, potentially allowing AI to generalize better to real-world scenarios and overcome the difficulties of using real data for training. Generally, synthetic data refer to artificially generated data that mimic the characteristics and structure of real data without being directly derived from actual observations (Fig. 1).

1.1 Why Synthetic Data?

Synthetic data are vital in AI research due to the challenges of acquiring real data, including time constraints, high costs, patient privacy concerns, and manual effort [122, 24, 31, 32]. They provide significant advantages by saving time and reducing the need for extensive manual annotation. The use of AI-generated content (AIGC) has proven effective across various domains, including medical imaging, where it serves both as a training resource for AI models and as a means of evaluating their performance with realistic yet hard-to-obtain data (see Table 1). Synthetic data offer precise control over properties such as shape, texture, and location, which is particularly valuable in medical applications. This control enables the creation of diverse and representative datasets for model training and provides useful examples for medical education and patient communication. Additionally, controllable synthetic

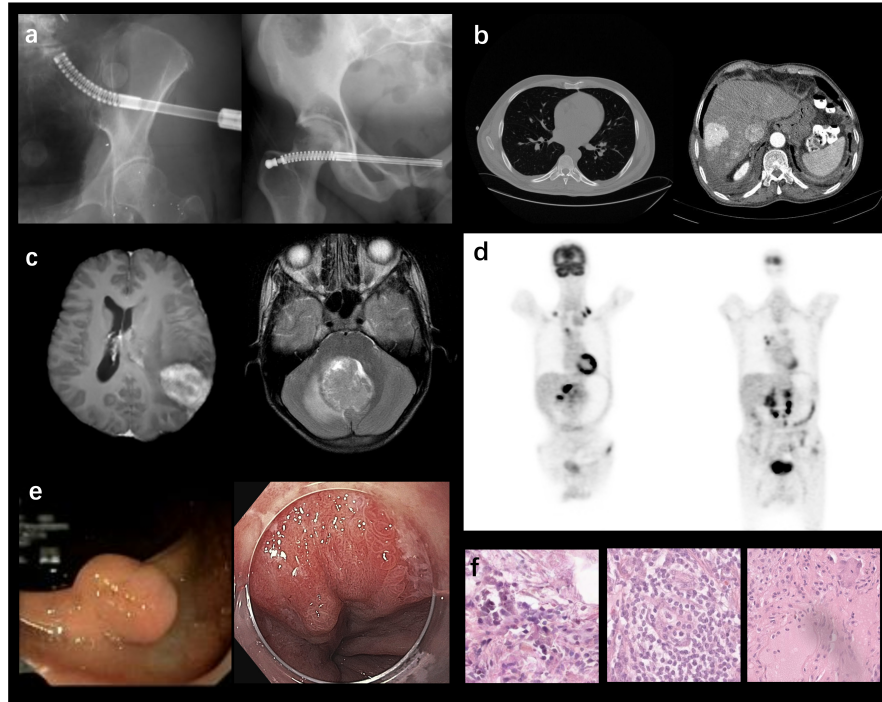


Fig. 1 Can you distinguish synthetic data from real data in different modalities? (a) X-ray image examples [26]. (b) CT image examples [33]. (c) MR image examples [77]. (d) PET image example. (e) Endoscopy image example [64]. (f) Histopathology image example [4].

tumors facilitate AI debugging and model diagnostics, enhancing the interpretability of AI behavior. Increasing evidence supports that synthetic data can improve AI performance, making it a powerful tool for advancing research and improving outcomes in fields like oncology.

1.2 Real vs. Synthetic Data

Unlike real data, which are collected from real-world imaging devices and represent true observations, synthetic data are created using algorithms, simulations, or models designed to replicate the properties of real data. We summarize the differences between real data and synthetic data as follows: *First*, real data are collected from actual imaging devices (e.g., X-ray, CT, MRI, PET, ultrasound, histopathology). Synthetic data are generated artificially using computational methods and simulations [42, 59]. *Second*, real data are often limited by practical constraints such as patient privacy, the cost of imaging, and the time required for data collection and annotation [13]. Synthetic data can be produced in large quantities without the ethical and logistical

Table 1 A summary of existing synthesis methods for medical imaging encompasses various aspects: body part, disease type, imaging modality, dataset utilized, and generative model type. Additionally, we are maintaining a [webpage](#) to track the latest publications and corresponding code repositories on data synthesis in medicine more comprehensively.

reference	body part	disease	modality	dataset	generative model
Teixeira et al. [97]	whole body	anomaly detection	X-ray	private dataset	GAN
Wu et al. [107]	chest	breast cancer	X-ray	DDSM [37]	GAN
Gao et al. [26]	bone & chest	COVID-19 lesion	X-ray	COVID-19 CXR [101]	GAN
Jin et al. [48]	chest	lung nodule	CT	LIDC [3]	GAN
Yao et al. [113]	chest	COVID-19	CT	LUNA16 [87]	hand-crafted designs
Jiang et al. [47]	chest	Covid-19	CT	COVID [112]	GAN
Jin et al. [49]	abdomen	liver & kidney tumors	CT	LiTS [8] & KiTS [38]	GAN
Wei et al. [105]	abdomen	pancreatic tumors	CT	private dataset	GAN
Lyu et al. [70]	abdomen	liver tumors	CT	LiTS [8]	GAN
Hu et al. [42]	abdomen	liver tumors	CT	LiTS [8]	hand-crafted designs
Li et al. [63]	abdomen	pancreatic tumors	CT	MSD [2]	hand-crafted designs
Chen et al. [9]	abdomen	tumors in liver, pancreas and kidney	CT	MSD [2] & KiTS [38]	Diffusion Model
Lai et al. [59]	abdomen	tumors in liver, pancreas and kidney	CT	MSD [2] & KiTS [38]	hand-crafted designs
Yu et al. [117]	brain	brain tumors	MRI	BraTS [72]	GAN
Han et al. [34]	brain	brain tumors	MRI	BraTS [72]	GAN
Zhao et al. [119]	abdomen	liver tumors	MRI	private dataset	GAN
Mukherjee et al. [73]	brain	brain tumors	MRI	BraTS [72]	GAN
Basaran et al. [6]	brain	brain tumors	MRI	WMH [54]	hand-crafted designs
Wang et al. [103]	brain	-	PET	private dataset	GAN
Luo et al. [69]	brain	-	PET	private dataset	GAN
Sharan et al. [88]	mitral valve	landmark detection	Endo.	surgical simulator [23]	GAN
Yoon et al. [116]	colon	polyp detection	Endo.	private dataset	GAN
Hou et al. [41]	tissue	cancer	Histo.	Kumar [55]	GAN
Xue et al. [110]	tissue	cancer	Histo.	PCam [100]	GAN
Aversa et al. [4]	tissue	cancer	Histo.	private dataset	Diffusion Model
Du et al. [19]	skin	dermatoscopic lesion	Dermo.	ISIC [17]	Diffusion Model

GAN stands for Generative Adversarial Network

Endo. refers to Endoscopy, Histo. refers to Histopathology, and Dermo. refers to dermoscopy.

issues associated with real data collection [10]. *Third*, real data require manual annotation by experts, which is time-consuming [78] and prone to human error [65]. Annotations of synthetic data can be automatically generated as part of the data creation process, ensuring consistency and accuracy [80].

2 Detecting Real Tumors in CT Scans

2.1 Tumors in Solid Organs

Solid tumors in organs like the liver, kidneys, and brain—such as hepatocellular carcinoma, renal cell carcinoma, and glioma—typically show well-defined margins and growth patterns [85], illustrated in Fig. 2. In CT images, early-stage tumors appear as small nodules with slightly blurred edges and homogeneous texture. As tumors advance, they grow larger, become irregular in shape, and exhibit significant mass

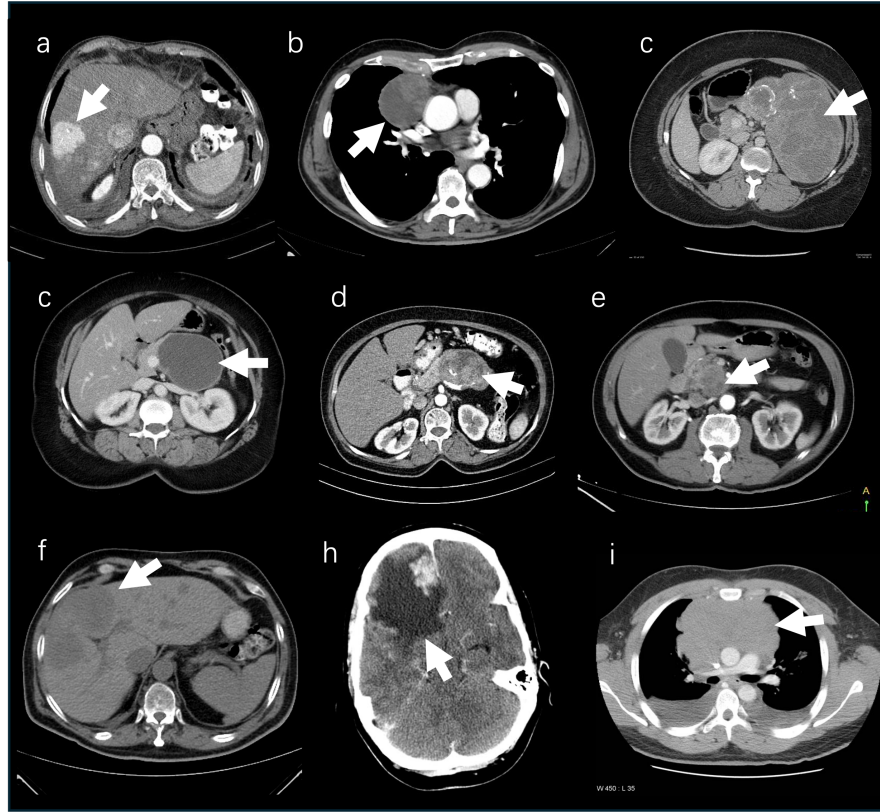


Fig. 2 Tumors in solid organs. (a) Hepatocellular carcinoma. (b) Thymoma. (c) Solid pseudopapillary tumor of the pancreas. (d) Pancreatic mucinous cystadenoma. (e) Pancreatic mucinous cystadenocarcinoma. (f) Pancreatic adenocarcinoma. (g) Neuroendocrine tumor in liver. (h) Meningioma. (i) Mediastinal lymphoma.

effect and infiltrative growth [20]. Advanced tumors may also show hemorrhage, necrosis, and fibrosis, leading to a heterogeneous appearance [89].

- **Liver tumors:** Hepatocellular carcinoma (HCC) is the most common malignant liver tumor. Early HCC presents as a small, well-differentiated nodule with a good prognosis and low metastatic potential [25]. CT imaging may show mass effect extending beyond the liver, displacement of blood vessels, intrahepatic venous thrombosis, and bile duct obstruction [82]. HCCs often exhibit intense arterial-phase enhancement due to neoangiogenesis and reduced portal triads, and typically appear hypoattenuating on venous phase scans [5].
- **Pancreatic tumors:** Pancreatic ductal adenocarcinoma (PDAC) constitutes the majority of malignant pancreatic tumors and is associated with a very poor prognosis and high morbidity. In the early stages, PDACs typically appear as homogeneous small nodules with blurred edges. Secondary findings associated

with advanced PDACs include contour abnormalities, abrupt termination of the biliary or pancreatic duct, pancreatic atrophy upstream from the mass, vascular encasement, etc. [57]. PDACs typically exhibit poor enhancement, appearing hypoattenuating relative to the surrounding pancreatic parenchyma. This hypoenhancement is attributed to the development of a dense fibroblastic stromal component in PDACs [22].

- **Kidney tumors:** Renal cell carcinoma (RCC) is the most common adult renal epithelial cancer, accounting for more than 90% of all renal malignancies [61]. The most prevalent subtype, clear cell RCC, presents as a homogeneously enhancing lesion during the corticomedullary phase and as a hypoattenuating renal lesion surrounded by homogeneously enhancing renal parenchyma in the nephrographic phase. Advanced clear cell RCC often appears heterogeneous in imaging due to the presence of hemorrhage, necrosis, and cysts, along with invasion into the renal pelvis, perirenal fat, or renal vessels [104].

2.2 Tumors in Tubular Organs

Tumors in tubular organs, illustrated in Fig. 3, have distinct growth patterns [74]: exophytic, where the tumor expands into the lumen, and invasive, where it penetrates the organ wall into adjacent structures. For instance, colon cancer progresses from stage 0, confined to the lining, to stage I, invading the submucosa, and stage II, extending through the wall without nearby invasion [21]. Stage III involves spread to lymph nodes, and stage IV features metastasis to distant organs such as the liver or lungs. We also highlight unique characteristics of representative tubular tumors to illustrate their specific behaviors and progression.

- **Esophageal tumors**, though rare, have a poor prognosis if malignant unless detected early and surgically removed [45]. Imaging studies include X-ray esophagography, CT, endoscopic ultrasound, and PET. Malignant strictures often show asymmetric narrowing with abrupt margins and irregular, nodular, or ulcerated surfaces. X-ray esophagography helps evaluate invasion of the muscularis mucosae for early-stage cancers. Key CT features are eccentric or circumferential wall thickening over 5 mm and periesophageal soft tissue and fat stranding [62, 94].
- **Stomach tumors**, primarily adenocarcinoma, are common and often asymptomatic when superficial. Up to 50% of patients may have nonspecific gastrointestinal symptoms like dyspepsia. Endoscopy is the most sensitive method for diagnosis, allowing direct visualization and biopsy. Initial detection is often through radiological methods, with CT imaging using negative contrast to reveal common features such as polypoid masses, wall thickening, or ulceration [93, 18, 62].
- **Colorectal tumors** are a leading gastrointestinal malignancy. Contrast-enhanced CT of the chest, abdomen, and pelvis is used for staging, detecting metastases, evaluating surgical options, and assessing treatment response. Colorectal cancers typically appear as soft tissue masses that narrow the bowel lumen. Larger tumors often show ulceration, mucinous tumors may appear as low-density masses

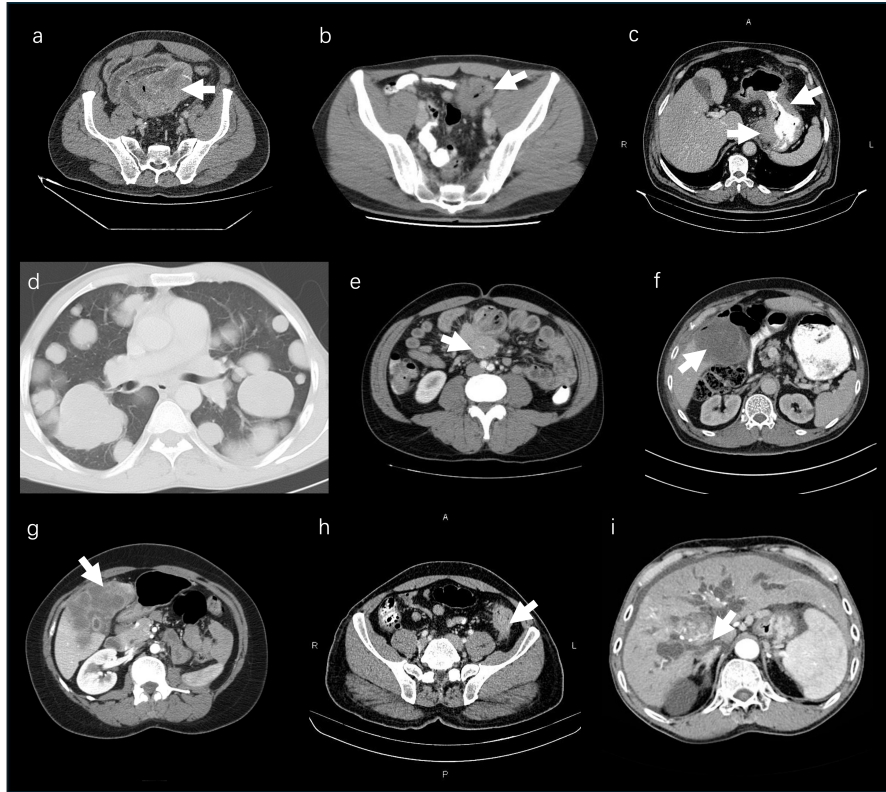


Fig. 3 Tumors in tubular organs. (a) Gastrointestinal stromal tumor. (b) Sigmoid colon cancer. (c) Gastric cancer. (d) Lung metastases. (e) Intestinal carcinoid tumor. (f) Gallbladder carcinoma. (g) Gallbladder adenocarcinoma. (h) Colon cancer. (i) Cholangiocarcinoma.

with low-density lymph nodes, and psammomatous calcifications can be seen in mucinous adenocarcinoma. [92, 30].

2.3 High Similarity in Early-Stage Tumors

Early-stage tumors (< 2cm) frequently exhibit similar imaging characteristics in CT scans, regardless of whether they originate in the liver, pancreas, or kidneys [12]. Should this finding be validated, it could carry profound implications for the application of generative AI in medical imaging. This implies that both modeling-based and learning-based approaches could be developed on a single tumor type with readily available annotated data and subsequently applied to synthesize various tumor types in other organs, for which data/annotation acquisition is more arduous.

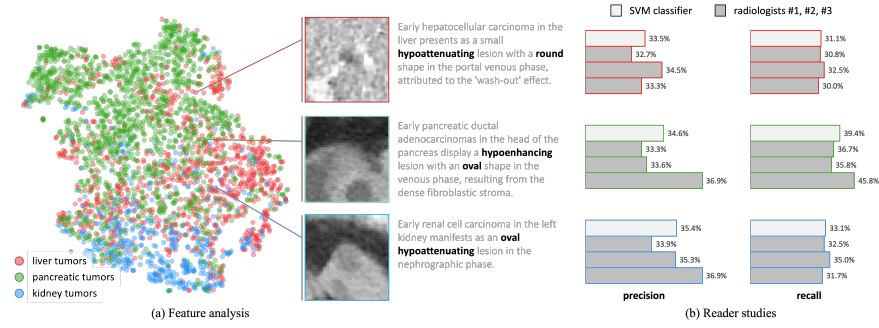


Fig. 4 *Feature analysis and reader study.* The left panel features a t-SNE (t-distributed stochastic neighbor embedding) visualization that maps the multidimensional Radiomics features of tumors from the liver, pancreas, and kidneys onto a two-dimensional space. This visualization underscores the substantial overlap in features among early-stage tumors from different organs, which may contribute to the challenges in correctly identifying their organ types. Complementing these findings, this study evaluates the efficacy of a support vector machine (SVM) classifier, which utilizes Radiomics Features [16, 102], in differentiating the organ types for the cropped tumors. The SVM classifier is trained to classify each tumor as originating from either the liver, pancreas, or kidneys—a three-way classification challenge. Parallel to the assessment of the SVM classifier, three expert radiologists conducted a similar evaluation by reviewing the original CT scans containing these tumors. The results displayed on the right panel reveal significant difficulties faced by both the SVM classifier and the radiologists when it comes to accurately pinpointing the origin of early-stage tumors. The precision and recall metrics for both the machine and human methods approximate the performance expected from random selection.

A study involving three expert radiologists was conducted to assess their ability to recognize the organ class of early-stage cancers [9]. Three expert radiologists, certified in accordance with the Quality Standards Act, participated in the reader study. Recognition results are presented in Fig. 4(b). The results were sufficiently compelling that medical professionals, boasting over five years of experience, could potentially mistake the synthetic tumors for genuine ones. The precision and recall scores, which approximate randomness, imply that the similarity in the appearance of early-stage tumors is such that even seasoned radiologists encounter difficulties when attempting to distinguish the organ types of these tumors.

The similarity of early-stage tumors can be evidenced by their Radiomics Feature¹ profiles [9]. From a *qualitative* standpoint, Fig. 4(a) depicts the feature mapping in a two-dimensional space via t-SNE. The appearance features of early-stage tumors manifest within a joint feature space, with no discernible segregation among different organ types. From a *quantitative* perspective, we trained a support vector machine (SVM) classifier to identify the organ types of early-stage tumors. To infer a general conclusion, we conducted ten repeated experiments and computed the precision and

¹ Utilization of the official radiomics feature repository [99] enables the extraction of appearance features, comprising 3D shape-based features, gray level co-occurrence matrix, gray level run length matrix, gray level size zone matrix, neighboring gray-tone difference matrix, and gray level dependence matrix.

recall metrics of the SVM classifier for both the training and test sets. The final results indicate that both the precision and recall metrics for the training set are close to 1.0, demonstrating that the SVM is effectively trained and has established a robust decision boundary for the training set. However, the precision scores for the test set approximate random chance, as illustrated in Fig. 4(b). This implies that even an effectively trained SVM classifier encounters difficulty in recognizing the organ types of unseen early-stage tumors.

3 Technical Barriers and Clinical Needs

3.1 Technical Barriers

AI development for real tumors faces key technical barriers: First, **data scarcity**: High-performance models need extensive annotated data, which is limited due to the time and expertise required for medical image and genomic annotation. Rare cancers further exacerbate this issue, leading to poor model performance on less common types [120, 121, 96, 14]. Second, **generalization to different organs**: AI models struggle to generalize across organs due to distinct anatomical structures and imaging modalities. Models trained on one organ, like the lung, perform poorly on others, such as the liver, due to differing tissue compositions and imaging techniques [118, 68]. Third, **generalization to different demographics**: Privacy laws restrict access to diverse datasets, impacting model robustness. Variations in imaging protocols and genetic differences across populations can lead to biased models that perform poorly on underrepresented groups [66].

3.2 Clinical Needs

Cancer research addresses critical clinical needs to improve patient outcomes and advance oncology. Key objectives include early cancer detection, developing effective treatments, and personalizing care strategies to enhance treatment success and system efficiency.

Early Detection and Diagnosis: Early detection and diagnosis of cancer are crucial for improving patient outcomes, as identifying cancer at an early stage often leads to more effective treatment and better survival rates. There is a critical need for screening methods with high sensitivity (ability to correctly identify those with cancer) and high specificity (ability to correctly identify those without cancer). Improved accuracy reduces false positives and false negatives, which are common issues in current screening practices. Besides, enhanced accuracy can help avoid overdiagnosis, where non-life-threatening cancers are treated unnecessarily, causing undue stress and potential harm to patients.

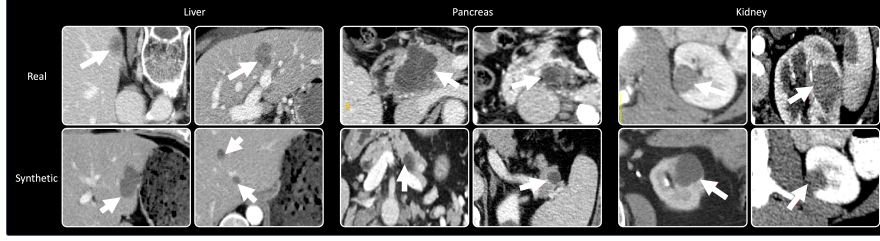


Fig. 5 Synthetic liver, pancreatic, and kidney tumors generated by Cellular Automata [59].

Health System Efficiency: For effective training in tumor detection across multiple organs, AI models traditionally require numerous annotated real tumor examples from each organ [123, 50, 79, 67]. However, these AI models often face difficulties in generalizing their ability to interpret images from different hospitals, a challenge compounded by varying imaging protocols, patient demographics, and scanner manufacturers [75, 111]. While the challenge of domain generalization could be alleviated if the AI is trained on a considerable number of annotated data from various domains [109, 114], it could take up to 25 human years for just annotating tumors in a specific organ [108, 1]. Collecting and annotating a comprehensive dataset that includes tumor examples from several organs (N) and numerous hospitals (M) is a formidable task, denoted by the complexity ($N \times M$). We hypothesize that *tumor synthesis could address this challenge by creating various tumor types across non-tumor images from multiple hospitals, even if only one type of tumor is available and annotated*. This approach can simplify the complexity from $N \times M$ to $1 \times M$.

Personalized Treatment Planning Tumors within the same type of cancer can vary significantly at the genetic and molecular levels. Personalized treatment planning requires comprehensive genomic profiling to identify specific mutations, gene expression patterns, and other molecular characteristics that drive an individual’s cancer. This information helps in selecting targeted therapies that are more likely to be effective for that particular tumor profile. Understanding the diversity of cancer cells within a tumor can inform treatment strategies that target multiple pathways and cell populations simultaneously.

4 Technology Trend I: Modeling-based Approaches

Hand-crafted tumor synthesis has been conducted in [43, 42, 44, 63]. This approach applies a sequence of hand-crafted morphological image-processing operations, including local selection, texture generation, shape generation, and post-processing, to generate realistic tumors for training AI models. The intrinsic observation about these operations is clinical knowledge. Taking liver tumors as an example, the mean HU intensity of hepatocellular carcinomas (tumors grown from liver cells) was 106 HU (with a range of 36-162 HU) [60]. Milder carcinomas usu-

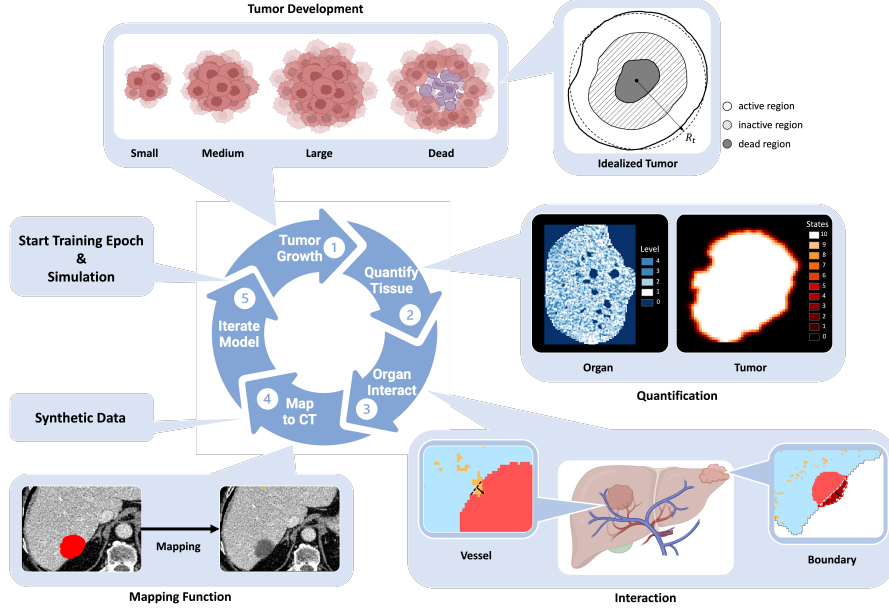


Fig. 6 ① Tumor development: The cellular automata simulate tumor growth from a single pixel to various sizes and even tumor death, producing synthetic tumors with diverse sizes, shapes, and textures. An idealized tumor is created to quantify development, with dead cells in the gray region, living quiescent cells in the inactive region, and proliferative cells in the active outer shell. ② Tissue quantification: The organ map in blue transforms CT images into distinct intensity levels affecting tumor development rate, while the red tumor map assigns values representing the tumor cell population. ③ Tumor interaction with boundaries and vessels: The tumor grows and exerts pressure against organ boundaries and deforms as it interacts with vessels. ④ Mapping synthetic tumors to CT images: A mapping function correlates the synthetic tumor with CT values, integrating the tumor's state with the original CT intensity. ⑤ Training segmentation models with synthetic data: Pixel2Cancer generates new synthetic data for each epoch to train the segmentation model.

ally lead to smaller, fewer spherical lesions, while multi-focal lesions (which means scattered small tumors) only appear in rare cases. Additionally, larger tumors usually display evident mass effects and are accompanied by capsule appearances that separate the tumor from the liver parenchyma [71]. This medical guidance, together with visual clues, determines the parameters and pipeline of this method.

Cellular Automata are computational models used to simulate complex systems through simple rules and interactions. They employ a grid of cells (pixels), where each cell is initially assigned a state between zero and ten to represent the tumor population. The basic element is the *cell*, which refers to a single *pixel* in the computed tomography (CT) image. Tumor growth and behavior are modeled based on specific rules that simulate processes such as proliferation, invasion, and death. These rules are derived from medical knowledge and are guided by an idealized tumor model that reflects real-world characteristics. The tumor state can then be integrated into the original CT images to generate synthetic tumors in different

organs. This tumor synthesis approach allows for sampling tumors at various stages and analyzing tumor-organ interactions. Motivated by this, Lai et al. [59] proposed Pixel2Cancer to simulate tumor growth (Fig. 6). Tumors generated by Pixel2Cancer are illustrated in Fig. 5.

4.1 Property of Pixel2Cancer

(i) Label-free. Pixel2Cancer can be applied as a label-free data synthesis approach, eliminating the need for manual per-voxel annotation. Previous learning-based approaches, such as GANs [28] and Diffusion models [40], are designed to learn the representation and distribution of tumors. While these approaches excel in generating natural images, synthesizing tumors in CT scans still requires significant amounts of paired tumor data. Moreover, when generating synthetic tumors, generative models also need masks to indicate the tumor locations and shapes [49], necessitating extensive manual efforts for training and synthesis.

(ii) Tumor development. Pixel2Cancer incorporates specific medical knowledge regarding tumor growth and appearance, enabling the simulation of realistic tumors. None of the existing synthetic approaches can adequately simulate tumor development in abdominal CT, and the primary challenges in current synthetic methods are the proliferation and invasion of tumors [36]. These processes in tumor growth are complex and interconnected, highly influenced by the surrounding environment [95, 106]. Consequently, synthetic tumors generated using current methods may conflict with normal organ structures and pose challenges when adapting them to different organs.

(iii) Early tumor detection & boundary segmentation. Early detection of small tumors is critical for timely cancer diagnosis. However, real datasets often lack sufficient instances due to the asymptomatic nature of early-stage patients. Pixel2Cancer can generate more small tumors to improve the sensitivity of segmentation models for small tumor detection. Additionally, Pixel2Cancer generates synthetic tumors with precise tumor masks, whereas real data annotations are often inaccurate at the boundaries, leading to *label noise* in boundary segmentation accuracy.

4.2 Clinical Perspectives

Tumors and genetic disorders from DNA mutations in single cells undergo complex growth processes [56]. Mutations during cell division lead to uncontrolled proliferation, forming neoplastic lesions that can be benign or malignant [27]. While both types follow similar growth principles, they differ in growth rate and invasiveness. Malignant tumors often grow rapidly, secreting growth factors or inducing surrounding stromal cells to do so, as seen in pancreatic IPMN lesions which grow larger and faster than benign ones [51]. Slow growth rates in renal tumors and hepatocellular

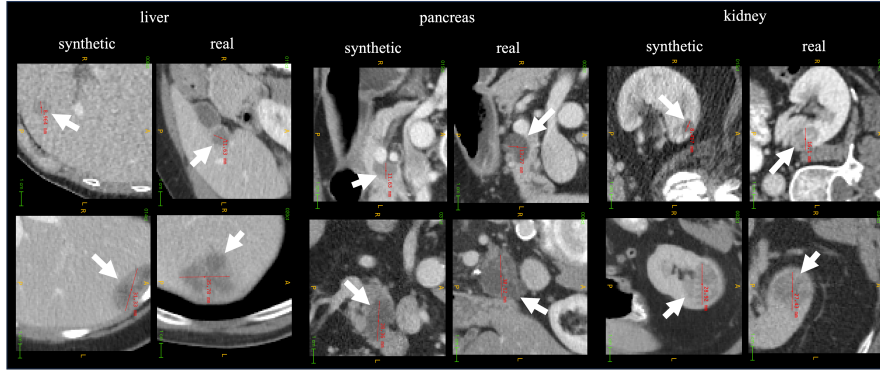


Fig. 7 Examples of synthetic tumors generated by DiffTumor [9] on the liver, pancreas, and kidney.

carcinoma correlate with lower malignancy [91]. Malignant tumors are invasive, gradually penetrating and destroying surrounding tissues, whereas benign tumors remain confined to their original sites. Even slowly growing malignant tumors can infiltrate neighboring structures. Tumor necrosis, caused by rapid proliferation exceeding vascular supply [39], appears as irregular hypo-attenuating areas in CT images and serves as a poor prognostic indicator [25]. The **death** rule models this necrosis. A hybrid cellular automaton model is proposed to simulate tumor development from single cells to invasive tumors, capturing their continuous progression and interactions within the microenvironment.

5 Technology Trend II: Learning-based Approaches

Generative Adversarial Networks (GANs) [29] have been extensively explored for generating synthetic tumors. Zhao et al. [119] introduced Tripartite-GAN, which simultaneously achieves contrast-enhanced magnetic resonance imaging synthesis and tumor detection. Mukherjee et al. [73] proposed AGGrGAN to generate synthetic MRI scans of brain tumors. However, the training process of GANs is often unstable, making it challenging to achieve convergence. Additionally, GANs can suffer from issues such as mode collapse, where the model generates limited diversity in outputs.

Diffusion Models [76, 52] provide more stable and reliable training compared to GANs, as they gradually denoise data, making the optimization process easier to control. Additionally, Diffusion Models are capable of generating high-quality, diverse samples with fewer issues related to mode collapse. DiffTumor [9] is the first to explore tumor synthesis in abdominal organs using Diffusion Models and demonstrates an efficient method for achieving generalizable tumor synthesis (Fig. 8). Tumors generated by DiffTumor are illustrated in Fig. 7.

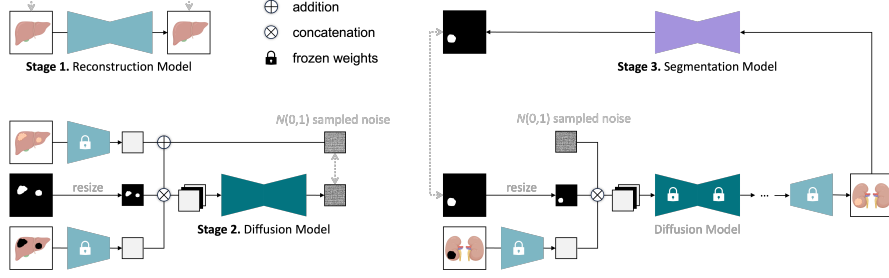


Fig. 8 Overview of DiffTumor framework. In pursuit of achieving generalizable tumor synthesis, DiffTumor encompasses three stages. ① The first stage is the training of an Autoencoder Model—comprising an encoder and a decoder—to learn comprehensive latent features. The learning objective in this stage entails image reconstruction conducted on 9,262 unlabeled three-dimensional CT volumes. Both the trained encoder and decoder are integral to subsequent stages. ② The second stage involves training a Diffusion Model—a specialized generative model—by utilizing latent features and tumor masks as conditions. Once trained, the model is capable of generating the requisite latent features for the reconstruction of CT volumes with tumors, utilizing arbitrary masks. ③ The third stage entails training a Segmentation Model with CT volumes of synthetic tumors, reconstructed by the decoder. Armed with a considerable repository of healthy CT volumes, DiffTumor has the capacity to generate an extensive collection of synthetic tumors, which vary in location, size, shape, texture, and intensity, thus contributing to the enhancement of AI models for tumor detection and segmentation.

5.1 Property of DiffTumor

(i) Reduced annotations for Diffusion Model. The quality of synthetic data produced by a generative model typically depends heavily on the quantity and diversity of the paired training data used during the training phase [11, 46]. Nevertheless, the relationship between the number of annotated real tumors required for training the Diffusion Model and the performance of the Segmentation Model has not been extensively studied. DiffTumor found that the relationship between the amount of paired training data and the quality of synthetic data is not necessarily linear. Remarkably, it requires only one annotated tumor to train the Diffusion Model and to generate synthetic tumors for the subsequent training of the Segmentation Model. This finding is in contrast to the conventional wisdom in computer vision [81], which often necessitates extensive datasets for training generative models. The results suggest that a smaller number of real tumors may suffice for training the Diffusion Model, particularly for early-stage tumors. Such a discovery could have profound implications for improving efficiency and reducing the costs associated with training generative models in the field of medical imaging.

(ii) Accelerated tumor synthesis. The speed at which synthetic tumors are generated is critical for the practical use of synthetic data, particularly for accelerating the training of segmentation models. The synthesis speed of DiffTumor is significantly affected by the choice of timestep (T). An investigation into the influence of timestep on the segmentation performance of the Segmentation Model has been conducted.

It is observed that using DDPM [40] with one-step sampling, DiffTumor cannot synthesize realistic textures for both organ and tumor, which negatively affects the training of the Segmentation Model. Conversely, by setting T higher than 1, DiffTumor can produce more realistic textures, leading to an enhanced performance of the segmentation model. Taking into account the balance between performance and synthesis efficiency, DiffTumor selects a timestep of $T = 4$ as the default setting for early tumor synthesis. This selection strikes a balance that allows for the generation of high-quality synthetic data while maintaining an acceptable level of efficiency.

5.2 Clinical Perspectives

This learning-based approach can be widely applied because of the similar growth dynamics observed in tumors across various types and locations. Tumorigenesis is a complex, multistage process involving cellular and histological transformations, from precancerous lesions to malignant tumors. This progression, driven by genetic mutations and functional changes known as the ‘hallmarks of cancer,’ is consistent across tumor types [58, 12, 35]. Early-stage tumors typically consist of well-to-moderately differentiated cells with mild atypia and invasiveness, showing rare hemorrhage and necrosis. They often appear as homogeneous nodules with slightly indistinct margins and small diameters in CT images [15, 90]. In contrast, advanced tumors exhibit significant infiltration and destruction of surrounding tissues, extending beyond the original site and potentially affecting adjacent structures [5]. As tumors become more malignant, rapid growth leads to ischemia and necrosis due to insufficient vascular supply, resulting in heterogeneous patterns in CT images with features like hemorrhage and fibrosis [115, 84]. These characteristics are consistent across different populations, ages, and genders.

6 Tumor Synthesis Benchmark

We evaluate the effectiveness of Pixel2Cancer and DiffTumor by comparing them with supervised models trained on real data and several prominent unsupervised anomaly segmentation methods. Table 2 highlights that synthetic data have significantly outperformed all these baseline methods, achieving a DSC of 59.77% and NSD of 61.29%. These results highlight the potential of synthetic strategies to avoid per-pixel manual annotation for tumor segmentation.

Table 2 *Comparison with state-of-the-art unsupervised methods.* We compare the initial label-free modeling-based methods with other unsupervised anomaly segmentation baselines, tumor synthesis strategies, and fully-supervised methods. Modeling-based methods significantly outperform all other state-of-the-art unsupervised baseline methods and even surpass the fully-supervised method with detailed *pixel-wise annotation*.

Liver tumor segmentation performance.

tumors	method	architecture	labeled / unlabeled CTs	DSC (%)	NSD (%)
none	PatchCore [83]	Resnet50	0 / 116	15.9	16.4
none	f-AnoGAN [86]	Customized [7]	0 / 116	19.0	16.9
none	VAE [53]	Customized [7]	0 / 116	24.6	23.6
synt	Yao et al. [113]	U-Net	0 / 116	32.8	31.3
real	fully-supervised	U-Net	101 / 0	56.7	58.0
synt	hand-crafted [42]	U-Net	0 / 116	59.8	61.3
synt	Pixel2Cancer [59]	U-Net	0 / 116	58.9	63.7
synt	DiffTumor [9]	U-Net	101 / 116	70.9	71.2

Pancreas tumor segmentation performance.

tumors	method	architecture	labeled / unlabeled CTs	DSC (%)	NSD (%)
real	fully-supervised	U-Net	96 / 0	57.5	56.5
synt	hand-crafted [42]	U-Net	0 / 104	54.1	52.2
synt	Pixel2Cancer [59]	U-Net	0 / 104	60.9	57.1
synt	DiffTumor [9]	U-Net	96 / 104	64.8	60.5

Kidney tumor segmentation performance.

tumors	method	architecture	labeled / unlabeled CTs	DSC (%)	NSD (%)
real	fully-supervised	U-Net	96 / 0	71.3	62.8
synt	hand-crafted [42]	U-Net	0 / 120	63.2	55.4
synt	Pixel2Cancer [59]	U-Net	0 / 120	73.2	65.0
synt	DiffTumor [9]	U-Net	96 / 120	84.2	76.6

6.1 Case Study: Fake Tumors, Real Results

Synthetic Liver Tumors. In synthetic liver tumors, synthetic data have demonstrated significant superiority over real data across all stages, from small to large tumors. The hand-crafted approach proposed by [42] outperforms real data, achieving a 2.3% improvement in DSC and a 3.3% improvement in NSD. Pixel2Cancer has shown superior performance in liver segmentation, with a DSC improvement of 2.2% and an NSD improvement of 5.7%. Additionally, DiffTumor has surpassed the performance of real data by 4.0% in DSC and 4.7% in NSD.

Synthetic Kidney Tumors. In kidney tumors, segmentation models have achieved superior performance when using synthetic data as data augmentation. Pixel2Cancer has shown superior performance in kidney segmentation, with a DSC improvement of 2.4% and an NSD improvement of 3.2%. Additionally, DiffTumor has surpassed the performance of real data by 7.0% in DSC and 6.7% in NSD.

Synthetic Pancreatic Tumors. Segmentation models also have achieved superior performance in pancreatic tumor segmentation when using synthetic data as data augmentation. Pixel2Cancer has shown superior performance in pancreas segmentation, with a DSC improvement of 3.9% and an NSD improvement of 1.9%. DiffTumor has surpassed the performance of real data by 8.2% in DSC and 9.4% in NSD.

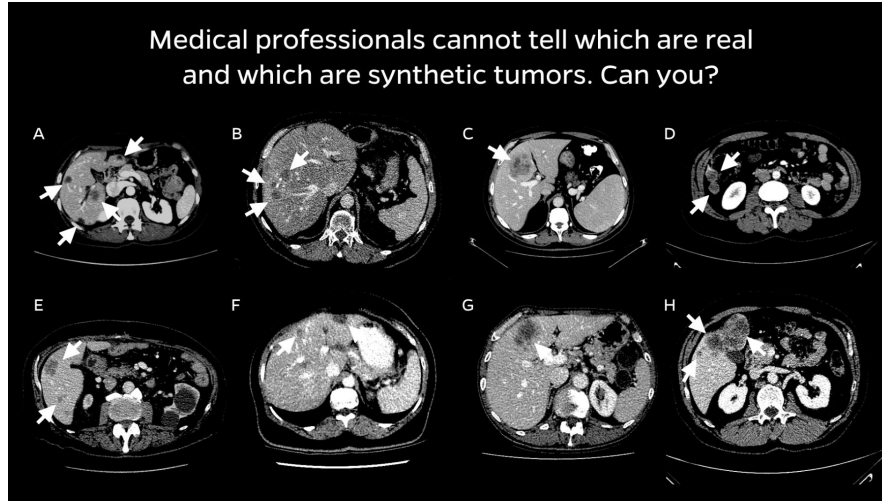


Fig. 9 Visual Turing Test. Can you find some examples of synthetic data in the CT images? Answers are in *Johns Hopkins Researchers Create Artificial Tumors to Help AI Detect Early-Stage Cancer*.

6.2 Visual Turing Test

The Turing Test, introduced by Alan Turing in “Computing Machinery and Intelligence” [98], assesses whether a machine can exhibit intelligent behavior indistinguishable from that of a human. We apply the Visual Turing Test to evaluate if synthetic tumors resemble real tumors. For this, we compared CT volumes containing real and synthetic tumors across various organs. Professionals, blinded to the origins of the samples, classified each volume as real or synthetic based on 3D views of continuous slice sequences (see Fig. 9).

Modeling-based approach (Pixel2Cancer). The outcome metrics, as presented in Table 3, unveil the performance evaluations by different radiologists. For the junior radiologist R1 (7 years of experience), metrics such as accuracy, sensitivity, and specificity all register below 40%. Notably, a specificity of 35.5% indicates that 64.5% of synthetic tumors are inaccurately identified as real. The intermediate radiologist R2 (9 years of experience) exhibits comparable metrics around 40%, with 59.2% of synthetic tumors causing confusion. Even the senior radiologist R3 (14 years of experience) misclassifies 44.1% of synthetic tumors as real, underscoring the formidable challenge posed even to seasoned professionals. These results emphasize the efficacy of our modeling-based approach (Pixel2Cancer) in achieving a realistic simulation of tumor development.

Learning-based approach (DiffTumor). Two professionals were involved in this test, with 4 and 11 years of experience, respectively. 60 CT volumes were arranged in random order and scrutinized independently by two professionals. The test outcomes are detailed in Table 3. Radiologists R1’s near-zero specificity scores suggested that synthetic data closely resembled real tumors, resulting in the misclassification of

Table 3 *Results of reader study.* Pixel2Cancer (top table): The test was conducted with three medical professionals having 7, 9, and 14 years of experience, respectively. Each professional evaluated 50 CT images for each organ, consisting of both real and synthetic tumors. They were tasked with categorizing each CT image as either *real*, *synthetic*, or *unsure*. DiffTumor (bottom table): Visual Turing test over three organs has been conducted with two radiologists (R1 and R2). Both radiologists are provided with 60 three-dimensional CT volumes of each organ, including 30 scans with real tumors and the remaining 30 with synthetic ones. Radiologists are tasked to label each CT volume as *real* or *synthetic*. A lower specificity score indicates a higher number of synthetic tumors being identified as real.

Modeling-based Approach: Pixel2Cancer

reader	metric	liver	pancreas	kidneys
R1	sensitivity (%)	100	95.0	95.5
	specificity (%)	27.3	22.7	26.7
	accuracy (%)	60.9	57.1	67.6
R2	sensitivity (%)	94.7	87.5	90.0
	specificity (%)	47.8	47.4	56.3
	accuracy (%)	69.1	65.7	75.0
R3	sensitivity (%)	100	100	100
	specificity (%)	45.4	55.6	57.9
	accuracy (%)	68.4	72.4	75.8

positives: real tumors ($N = 25$); negatives: synthetic tumors ($N = 25$).

Learning-based Approach: DiffTumor

reader	metric	liver	pancreas	kidneys
R1	sensitivity (%)	100	97.1	92.9
	specificity (%)	2.9	0.0	3.1
	accuracy (%)	45.0	56.7	45.0
R2	sensitivity (%)	84.6	100	100
	specificity (%)	47.1	44.0	65.6
	accuracy (%)	63.3	76.7	81.7

positives: real tumors ($N = 30$); negatives: synthetic tumors ($N = 30$).

most synthetic tumors as real. Consequently, R1’s accuracy scores hovered around 50%. Conversely, R2, with more experience, exhibited higher specificity scores compared to R1, approaching 50%. This implies that nearly 50% of synthetic samples were correctly identified as synthetic, indicating a better discernment between real and synthetic tumors by R2. These findings affirm the effectiveness of DiffTumor in generating visually realistic tumors.

7 Conclusion

In this chapter, we delved into the concept of synthetic data and its application in medical fields, with a particular focus on cancer research. We defined synthetic data and discussed its critical role in cancer research, such as improving data diversity, protecting patient privacy, and enabling robust research in tumor detection, diagnosis, and treatment. We explored the challenges and opportunities presented in cancer research. Despite these challenges, synthetic data offers significant opportunities,

such as enhancing the training of machine learning models, supporting large-scale studies without privacy concerns, and fostering innovation in personalized medicine. Additionally, we highlighted promising approaches and future directions for the use of synthetic data in cancer research, including modeling-based methods and learning-based methods. These methods are opening new avenues for more accurate and comprehensive cancer research, enabling researchers to simulate various scenarios and treatments, and ultimately contributing to better patient outcomes.

Acknowledgement. This work was supported by the Lustgarten Foundation for Pancreatic Cancer Research and the Patrick J. McGovern Foundation Award.

References

1. Abi Nader, C., Vetil, R., Wood, L.K., Rohe, M.M., Bône, A., Karteszi, H., Vullierme, M.P.: Automatic detection of pancreatic lesions and main pancreatic duct dilatation on portal venous ct scans using deep learning. *Investigative Radiology* (2023)
2. Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., et al.: The medical segmentation decathlon. *Nature communications* **13**(1), 1–13 (2022)
3. Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al.: The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics* **38**(2), 915–931 (2011)
4. Aversa, M., Nobis, G., Hägele, M., Standvoss, K., Chirica, M., Murray-Smith, R., Alaa, A.M., Ruff, L., Ivanova, D., Samek, W., et al.: Diffinfinite: Large mask-image synthesis via parallel random patch diffusion in histopathology. *Advances in Neural Information Processing Systems* **36** (2024)
5. Ayuso, C., Rimola, J., Vilana, R., Burrel, M., Darnell, A., García-Criado, Á., Bianchi, L., Belmonte, E., Caparroz, C., Barrufet, M., et al.: Diagnosis and staging of hepatocellular carcinoma (hcc): current guidelines. *European journal of radiology* **101**, 72–81 (2018)
6. Basaran, B.D., Zhang, W., Qiao, M., Kainz, B., Matthews, P.M., Bai, W.: Lesionmix: A lesion-level data augmentation method for medical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 73–83. Springer (2023)
7. Baur, C., Denner, S., Wiestler, B., Navab, N., Albarqouni, S.: Autoencoders for unsupervised anomaly segmentation in brain mr images: a comparative study. *Medical Image Analysis* **69**, 101952 (2021)
8. Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G., et al.: The liver tumor segmentation benchmark (lits). *Medical Image Analysis* **84**, 102680 (2023)
9. Chen, Q., Chen, X., Song, H., Xiong, Z., Yuille, A., Wei, C., Zhou, Z.: Towards generalizable tumor synthesis. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024). URL <https://github.com/MrGiovanni/DiffTumor>
10. Chiruvella, V., Guddati, A.K., et al.: Ethical issues in patient data ownership. *Interactive journal of medical research* **10**(2), e22269 (2021)
11. Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., Haworth, A.: A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology* **65**(5), 545–563 (2021)
12. Choi, J.Y., Lee, J.M., Sirlin, C.B.: Ct and mr imaging diagnosis and staging of hepatocellular carcinoma: part i. development, growth, and spread: key pathologic and imaging aspects. *Radiology* **272**(3), 635–654 (2014)

13. Chou, Y.C., Li, B., Fan, D.P., Yuille, A., Zhou, Z.: Acquiring weak annotations for tumor localization in temporal and volumetric data. *Machine Intelligence Research* pp. 1–13 (2024). URL <https://github.com/johnson111788/Drag-Drop>
14. Chou, Y.C., Zhou, Z., Yuille, A.: Embracing massive medical data. *arXiv preprint arXiv:2407.04687* (2024). URL <https://github.com/MrGiovanni/OnlineLearning>
15. Chu, L.C., Goggins, M.G., Fishman, E.K.: Diagnosis and detection of pancreatic cancer. *The Cancer Journal* **23**(6), 333–342 (2017)
16. Chu, L.C., Park, S., Kawamoto, S., Fouladi, D.F., Shayesteh, S., Zinreich, E.S., Graves, J.S., Horton, K.M., Hruban, R.H., Yuille, A.L., et al.: Utility of ct radiomics features in differentiation of pancreatic ductal adenocarcinoma from normal pancreatic tissue. *American Journal of Roentgenology* **213**(2), 349–357 (2019)
17. Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al.: Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, pp. 168–172. IEEE (2018)
18. Davis, G.B., Blanchard, D.K., Hatch III, G.F., Wertheimer-Hatch, L., Hatch, K.F., Foster Jr, R.S., Skandalakis, J.E.: Tumors of the stomach. *World journal of surgery* **24**(4), 412–420 (2000)
19. Du, S., Wang, X., Lu, Y., Zhou, Y., Zhang, S., Yuille, A., Li, K., Zhou, Z.: Boosting dermatoscopic lesion segmentation via diffusion models with visual and textual prompts. *arXiv preprint arXiv:2310.02906* (2023)
20. Dunnick, N.R.: Renal cell carcinoma: staging and surveillance. *Abdominal Radiology* **41**, 1079–1085 (2016)
21. Edge, S.B., Compton, C.C.: The american joint committee on cancer: the 7th edition of the ajcc cancer staging manual and the future of tm. *Annals of surgical oncology* **17**(6), 1471–1474 (2010)
22. Elbanna, K.Y., Jang, H.J., Kim, T.K.: Imaging diagnosis and staging of pancreatic ductal adenocarcinoma: a comprehensive review. *Insights into imaging* **11**(1), 1–13 (2020)
23. Engelhardt, S., Sauerzapf, S., Preim, B., Karck, M., Wolf, I., De Simone, R.: Flexible and comprehensive patient-specific mitral valve silicone models with chordae tendineae made from 3d-printable molds. *International journal of computer assisted radiology and surgery* **14**, 1177–1186 (2019)
24. Feng, R., Zhou, Z., Gotway, M.B., Liang, J.: Parts2whole: Self-supervised contrastive learning via reconstruction. In: *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*, pp. 85–95. Springer (2020)
25. Fowler, K.J., Burgoyne, A., Fraum, T.J., Hosseini, M., Ichikawa, S., Kim, S., Kitao, A., Lee, J.M., Paradis, V., Taouli, B., et al.: Pathologic, molecular, and prognostic radiologic features of hepatocellular carcinoma. *Radiographics* **41**(6), 1611–1631 (2021)
26. Gao, C., Killeen, B.D., Hu, Y., Grupp, R.B., Taylor, R.H., Armand, M., Unberath, M.: Synthetic data accelerates the development of generalizable learning-based algorithms for x-ray image analysis. *Nature Machine Intelligence* **5**(3), 294–308 (2023)
27. Golias, C., Charalabopoulos, A., Charalabopoulos, K.: Cell proliferation and cell cycle control: a mini review. *International journal of clinical practice* **58**(12), 1134–1141 (2004)
28. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. *Advances in neural information processing systems* **27** (2014)
29. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. *Communications of the ACM* **63**(11), 139–144 (2020)
30. Griswold, D., Corbett, T.H.: A colon tumor model for anticancer agent evaluation. *Cancer* **36**(S6), 2441–2444 (1975)
31. Haghighi, F., Hosseinzadeh Taher, M.R., Zhou, Z., Gotway, M.B., Liang, J.: Learning semantics-enriched representation via self-discovery, self-classification, and self-restoration.

- In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 137–147. Springer (2020). URL <https://github.com/fhaghighi/SemanticGenesis>
32. Haghighi, F., Taher, M.R.H., Zhou, Z., Gotway, M.B., Liang, J.: Transferable visual words: Exploiting the semantics of anatomical patterns for self-supervised learning. *IEEE Transactions on Medical Imaging* (2021). URL <https://github.com/fhaghighi/SemanticGenesis>
 33. Hamamci, I.E., Er, S., Simsar, E., Tezcan, A., Simsek, A.G., Almas, F., Esirgun, S.N., Reynaud, H., Pati, S., Bluethgen, C., et al.: Generatect: text-guided 3d chest ct generation. In: *Proceedings of the European Conference on Computer Vision (ECCV)* (2024)
 34. Han, C., Rundo, L., Araki, R., Furukawa, Y., Mauri, G., Nakayama, H., Hayashi, H.: Infinite brain mr images: Prgan-based data augmentation for tumor detection. In: *Neural approaches to dynamics of signal exchanges*, pp. 291–303. Springer (2019)
 35. Hanahan, D., Weinberg, R.A.: The hallmarks of cancer. *cell* **100**(1), 57–70 (2000)
 36. Harpold, H.L., Alvord Jr, E.C., Swanson, K.R.: The evolution of mathematical modeling of glioma proliferation and invasion. *Journal of Neuropathology & Experimental Neurology* **66**(1), 1–9 (2007)
 37. Heath, M., Bowyer, K., Kopans, D., Kegelmeyer Jr, P., Moore, R., Chang, K., Munishkumar, S.: Current status of the digital database for screening mammography. In: *Digital Mammography: Nijmegen, 1998*, pp. 457–460. Springer (1998)
 38. Heller, N., Isensee, F., Maier-Hein, K.H., Hou, X., Xie, C., Li, F., Nan, Y., Mu, G., Lin, Z., Han, M., et al.: The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. *Medical image analysis* **67**, 101821 (2021)
 39. Hiraoka, N., Ino, Y., Sekine, S., Tsuda, H., Shimada, K., Kosuge, T., Zavada, J., Yoshida, M., Yamada, K., Koyama, T., et al.: Tumour necrosis is a postoperative prognostic marker for pancreatic cancer patients with a high interobserver reproducibility in histological evaluation. *British journal of cancer* **103**(7), 1057–1065 (2010)
 40. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020)
 41. Hou, L., Agarwal, A., Samaras, D., Kurc, T.M., Gupta, R.R., Saltz, J.H.: Robust histopathology image analysis: To label or to synthesize? In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8533–8542 (2019)
 42. Hu, Q., Chen, Y., Xiao, J., Sun, S., Chen, J., Yuille, A.L., Zhou, Z.: Label-free liver tumor segmentation. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7422–7432 (2023). URL <https://github.com/MrGiovanni/SyntheticTumors>
 43. Hu, Q., Xiao, J., Chen, Y., Sun, S., Chen, J.N., Yuille, A., Zhou, Z.: Synthetic tumors make ai segment tumors better. *NeurIPS Workshop on Medical Imaging meets NeurIPS* (2022). URL <https://github.com/MrGiovanni/SyntheticTumors>
 44. Hu, Q., Yuille, A., Zhou, Z.: Synthetic data as validation. *arXiv preprint arXiv:2310.16052* (2023). URL <https://github.com/MrGiovanni/SyntheticValidation>
 45. Iyer, R., Dubrow, R.: Imaging of esophageal cancer. *Cancer Imaging* **4**(2), 125 (2004)
 46. Jaipuria, N., Zhang, X., Bhasin, R., Arafa, M., Chakravarty, P., Shrivastava, S., Manglani, S., Murali, V.N.: Deflating dataset bias using synthetic data augmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 772–773 (2020)
 47. Jiang, Y., Chen, H., Loew, M., Ko, H.: Covid-19 ct image synthesis with a conditional generative adversarial network. *IEEE Journal of Biomedical and Health Informatics* **25**(2), 441–452 (2020)
 48. Jin, D., Xu, Z., Tang, Y., Harrison, A.P., Mollura, D.J.: Ct-realistic lung nodule simulation from 3d conditional generative adversarial networks for robust lung segmentation. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 732–740. Springer (2018)
 49. Jin, Q., Cui, H., Sun, C., Meng, Z., Su, R.: Free-form tumor synthesis in computed tomography images via richer generative adversarial network. *Knowledge-Based Systems* **218**, 106753 (2021)

50. Kang, M., Li, B., Zhu, Z., Lu, Y., Fishman, E.K., Yuille, A., Zhou, Z.: Label-assemble: Leveraging multiple datasets with partial labels. In: IEEE International Symposium on Biomedical Imaging, pp. 1–5. IEEE (2023). URL <https://github.com/MrGiovanni/LabelAssemble>
51. Kang, M.J., Jang, J.Y., Kim, S.J., Lee, K.B., Ryu, J.K., Kim, Y.T., Yoon, Y.B., Kim, S.W.: Cyst growth rate predicts malignancy in patients with branch duct intraductal papillary mucinous neoplasms. *Clinical Gastroenterology and Hepatology* **9**(1), 87–93 (2011)
52. Khader, F., Müller-Franzes, G., Tayebi Arasteh, S., Han, T., Haarbuerger, C., Schulze-Hagen, M., Schad, P., Engelhardt, S., Baeßler, B., Foersch, S., et al.: Denoising diffusion probabilistic models for 3d medical image generation. *Scientific Reports* **13**(1), 7303 (2023)
53. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)
54. Kuijff, H.J., Biesbroek, J.M., De Bresser, J., Heinen, R., Andermatt, S., Bento, M., Berseth, M., Belyaev, M., Cardoso, M.J., Casamitjana, A., et al.: Standardized assessment of automatic segmentation of white matter hyperintensities and results of the wmh segmentation challenge. *IEEE transactions on medical imaging* **38**(11), 2556–2568 (2019)
55. Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., Sethi, A.: A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging* **36**(7), 1550–1560 (2017)
56. Kumar, V., Abbas, A., Aster, J.C.: Robbins basic pathology. Elsevier Health Sciences (2017)
57. Laeseke, P.F., Chen, R., Jeffrey, R.B., Brentnall, T.A., Willmann, J.K.: Combining in vitro diagnostics with in vivo imaging for earlier detection of pancreatic ductal adenocarcinoma: challenges and solutions. *Radiology* **277**(3), 644–661 (2015)
58. Lahouel, K., Younes, L., Danilova, L., Giardiello, F.M., Hruban, R.H., Groopman, J., Kinzler, K.W., Vogelstein, B., Geman, D., Tomasetti, C.: Revisiting the tumorigenesis timeline with a data-driven generative model. *Proceedings of the National Academy of Sciences* **117**(2), 857–864 (2020)
59. Lai, Y., Chen, X., Wang, A., Yuille, A., Zhou, Z.: From pixel to cancer: Cellular automata in computed tomography. arXiv preprint arXiv:2403.06459 (2024). URL <https://github.com/MrGiovanni/Pixel2Cancer>
60. Lee, K., O'Malley, M., Haider, M., Hanbidge, A.: Triple-phase mdct of hepatocellular carcinoma. *American Journal of Roentgenology* **182**(3), 643–649 (2004)
61. Leveridge, M.J., Bostrom, P.J., Koulouris, G., Finelli, A., Lawrentschuk, N.: Imaging renal cell carcinoma with ultrasonography, ct and mri. *Nature Reviews Urology* **7**(6), 311–325 (2010)
62. Lewin, K.J., Appelman, H.D., et al.: Tumors of the esophagus and stomach. American Registry of Pathology (1996)
63. Li, B., Chou, Y.C., Sun, S., Qiao, H., Yuille, A., Zhou, Z.: Early detection and localization of pancreatic cancer by label-free tumor synthesis. MICCAI Workshop on Big Task Small Data, 1001-AI (2023). URL <https://github.com/MrGiovanni/SyntheticTumors>
64. Li, C., Liu, H., Liu, Y., Feng, B.Y., Li, W., Liu, X., Chen, Z., Shao, J., Yuan, Y.: Endora: Video generation models as endoscopy simulators. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (2024)
65. Li, W., Qu, C., Chen, X., Bassi, P.R., Shi, Y., Lai, Y., Yu, Q., Xue, H., Chen, Y., Lin, X., et al.: Abdomenatlas: A large-scale, detailed-annotated, & multi-center dataset for efficient transfer learning and open algorithmic benchmarking. *Medical Image Analysis* p. 103285 (2024). URL <https://github.com/MrGiovanni/AbdomenAtlas>
66. Li, W., Yuille, A., Zhou, Z.: How well do supervised models transfer to 3d image segmentation? In: International Conference on Learning Representations (2024). URL <https://github.com/MrGiovanni/SuPreM>
67. Liu, J., Zhang, Y., Chen, J.N., Xiao, J., Lu, Y., A Landman, B., Yuan, Y., Yuille, A., Tang, Y., Zhou, Z.: Clip-driven universal model for organ segmentation and tumor detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 21152–21164 (2023). URL <https://github.com/ljwztc/CLIP-Driven-Universal-Model>

68. Liu, J., Zhang, Y., Wang, K., Yavuz, M.C., Chen, X., Yuan, Y., Li, H., Yang, Y., Yuille, A., Tang, Y., et al.: Universal and extensible language-vision models for organ segmentation and tumor detection from abdominal computed tomography. *Medical Image Analysis* p. 103226 (2024). URL <https://github.com/ljwztc/CLIP-Driven-Universal-Model>
69. Luo, Y., Zhou, L., Zhan, B., Fei, Y., Zhou, J., Wang, Y., Shen, D.: Adaptive rectification based adversarial network with spectrum constraint for high-quality pet image synthesis. *Medical Image Analysis* **77**, 102335 (2022)
70. Lyu, F., Ye, M., Ma, A.J., Yip, T.C.F., Wong, G.L.H., Yuen, P.C.: Learning from synthetic ct images via test-time training for liver tumor segmentation. *IEEE transactions on medical imaging* **41**(9), 2510–2520 (2022)
71. M. Cunha, G., Fowler, K.J., Roudenko, A., Taouli, B., Fung, A.W., Elsayes, K.M., Marks, R.M., Cruite, I., Horvat, N., Chernyak, V., et al.: How to use li-rads to report liver ct and mri observations. *RadioGraphics* **41**(5), 1352–1367 (2021)
72. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging* **34**(10), 1993–2024 (2014)
73. Mukherjee, D., Saha, P., Kaplun, D., Sinitca, A., Sarkar, R.: Brain tumor image generation using an aggregation of gan models with style transfer. *Scientific Reports* **12**(1), 9141 (2022)
74. Nerad, E., Lahaye, M.J., Maas, M., Nelemans, P., Bakers, F.C., Beets, G.L., Beets-Tan, R.G.: Diagnostic accuracy of ct for local staging of colon cancer: a systematic review and meta-analysis. *American Journal of Roentgenology* **207**(5), 984–995 (2016)
75. Orbes-Arteaga, M., Varsavsky, T., Sudre, C.H., Eaton-Rosen, Z., Haddow, L.J., Sørensen, L., Nielsen, M., Pai, A., Ourselin, S., Modat, M., et al.: Multi-domain adaptation in brain mri through paired consistency and adversarial learning. In: *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, pp. 54–62. Springer (2019)
76. Özbey, M., Dalmaz, O., Dar, S.U., Bedel, H.A., Öztürk, Ş., Güngör, A., Çukur, T.: Unsupervised medical image translation with adversarial diffusion models. *IEEE Transactions on Medical Imaging* (2023)
77. Park, J.E., Eun, D., Kim, H.S., Lee, D.H., Jang, R.W., Kim, N.: Generative adversarial network for glioblastoma ensures morphologic variations and improves diagnostic model for isocitrate dehydrogenase mutant type. *Scientific reports* **11**(1), 9912 (2021)
78. Park, S., Chu, L., Fishman, E., Yuille, A., Vogelstein, B., Kinzler, K., Horton, K., Hruban, R., Zinreich, E., Fouladi, D.F., et al.: Annotated normal ct data of the abdomen for deep learning: Challenges and strategies for implementation. *Diagnostic and interventional imaging* **101**(1), 35–44 (2020)
79. Qu, C., Zhang, T., Qiao, H., Liu, J., Tang, Y., Yuille, A., Zhou, Z.: Abdomenatlas-8k: Annotating 8,000 abdominal ct volumes for multi-organ segmentation in three weeks. *Conference on Neural Information Processing Systems* (2023). URL <https://github.com/MrGiovanni/AbdomenAtlas>
80. Qu, C., Zhang, T., Qiao, H., Tang, Y., Yuille, A.L., Zhou, Z., et al.: Abdomenatlas-8k: Annotating 8,000 ct volumes for multi-organ segmentation in three weeks. *Advances in Neural Information Processing Systems* **36** (2024)
81. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* **1**(2), 3 (2022)
82. Reynolds, A.R., Furlan, A., Fetzer, D.T., Sasatomi, E., Borhani, A.A., Heller, M.T., Tublin, M.E.: Infiltrative hepatocellular carcinoma: what radiologists need to know. *Radiographics* **35**(2), 371–386 (2015)
83. Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., Gehler, P.: Towards total recall in industrial anomaly detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14318–14328 (2022)
84. Saar, B., Kellner-Weldon, F.: Radiological diagnosis of hepatocellular carcinoma. *Liver International* **28**(2), 189–199 (2008)
85. Sahani, D.V., Samir, A.E.: *Abdominal Imaging E-Book: Expert Radiology Series*. Elsevier Health Sciences (2016)

86. Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis* (2019)
87. Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., et al.: Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. *Medical image analysis* **42**, 1–13 (2017)
88. Sharan, L., Romano, G., Koehler, S., Kelm, H., Karck, M., De Simone, R., Engelhardt, S.: Mutually improved endoscopic image synthesis and landmark detection in unpaired image-to-image translation. *IEEE Journal of Biomedical and Health Informatics* **26**(1), 127–138 (2021)
89. Silverman, P.M.: *Oncologic imaging: a multidisciplinary approach*. Elsevier Health Sciences (2012)
90. Skarin, A.T.: *Atlas of Diagnostic Oncology E-Book*. Elsevier Health Sciences (2015)
91. Smaldone, M.C., Kutikov, A., Eggleston, B.L., Canter, D.J., Viterbo, R., Chen, D.Y., Jewett, M.A., Greenberg, R.E., Uzzo, R.G.: Small renal masses progressing to metastases under active surveillance: a systematic review and pooled analysis. *Cancer* **118**(4), 997–1006 (2012)
92. Soga, J.: Early-stage carcinoids of the gastrointestinal tract: an analysis of 1914 reported cases. *Cancer: Interdisciplinary International Journal of the American Cancer Society* **103**(8), 1587–1595 (2005)
93. Stout, A.P.: *Tumors of the stomach*. Armed Forces Institute of Pathology (1953)
94. Stout, A.P., Lattes, R.: *Tumors of the esophagus*. Armed Forces Institute of Pathology (1957)
95. Tanase, M., Waliszewski, P.: On complexity and homogeneity measures in predicting biological aggressiveness of prostate cancer; implication of the cellular automata model of tumor growth. *Journal of surgical oncology* **112**(8), 791–801 (2015)
96. Tang, Y., Liu, J., Zhou, Z., Yu, X., Huo, Y.: Efficient 3d representation learning for medical image analysis. *World Scientific Annual Review of Artificial Intelligence* (2024)
97. Teixeira, B., Singh, V., Chen, T., Ma, K., Tamersoy, B., Wu, Y., Balashova, E., Comaniciu, D.: Generating synthetic x-ray images of a person from the surface geometry. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9059–9067 (2018)
98. Turing, A.M.: *Computing machinery and intelligence*. Springer (2009)
99. Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.C., Pieper, S., Aerts, H.J.: Computational radiomics system to decode the radiographic phenotype. *Cancer research* **77**(21), e104–e107 (2017)
100. Veeling, B.S., Linmans, J., Winkens, J., Cohen, T., Welling, M.: Rotation equivariant cnns for digital pathology. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 210–218. Springer (2018)
101. Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., Pinheiro, P.R.: Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. *Ieee Access* **8**, 91916–91923 (2020)
102. Wang, H., Zhou, Z., Li, Y., Chen, Z., Lu, P., Wang, W., Liu, W., Yu, L.: Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18f-fdg pet/ct images. *EJNMMI research* **7**(1), 1–11 (2017)
103. Wang, Y., Zhou, L., Wang, L., Yu, B., Zu, C., Lalush, D.S., Lin, W., Wu, X., Zhou, J., Shen, D.: Locality adaptive multi-modality gans for high-quality pet image synthesis. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 329–337. Springer (2018)
104. Wang, Z.J., Westphalen, A.C., Zagoria, R.J.: CT and MRI of small renal masses. *Br J Radiol* **91**(1087), 20180131 (2018)
105. Wei, Z., Chen, Y., Guan, Q., Hu, H., Zhou, Q., Li, Z., Xu, X., Frangi, A., Chen, F.: Pancreatic image augmentation based on local region texture synthesis for tumor segmentation. In: *International Conference on Artificial Neural Networks*, pp. 419–431. Springer (2022)
106. Wong, K.C., Summers, R.M., Kebebew, E., Yao, J.: Pancreatic tumor growth prediction with multiplicative growth and image-derived motion. In: *International Conference on Information Processing in Medical Imaging*, pp. 501–513. Springer (2015)

107. Wu, E., Wu, K., Cox, D., Lotter, W.: Conditional infilling gans for data augmentation in mammogram classification. In: Image Analysis for Moving Organ, Breast, and Thoracic Images: Third International Workshop, pp. 98–106. Springer (2018)
108. Xia, Y., Yu, Q., Chu, L., Kawamoto, S., Park, S., Liu, F., Chen, J., Zhu, Z., Li, B., Zhou, Z., et al.: The felix project: Deep networks to detect pancreatic neoplasms. medRxiv (2022)
109. Xiao, J., Yu, L., Zhou, Z., Bai, Y., Xing, L., Yuille, A., Zhou, Y.: Catenorm: Categorical normalization for robust medical image segmentation. In: MICCAI Workshop on Domain Adaptation and Representation Transfer, pp. 129–146. Springer (2022). URL <https://github.com/lambert-x/CateNorm>
110. Xue, Y., Ye, J., Zhou, Q., Long, L.R., Antani, S., Xue, Z., Cornwell, C., Zaino, R., Cheng, K.C., Huang, X.: Selective synthetic augmentation with histogan for improved histopathology image classification. Medical image analysis **67**, 101816 (2021)
111. Yan, W., Huang, L., Xia, L., Gu, S., Yan, F., Wang, Y., Tao, Q.: Mri manufacturer shift and adaptation: increasing the generalizability of deep learning segmentation for mr images acquired with different scanners. Radiology: Artificial Intelligence **2**(4), e190195 (2020)
112. Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., Xie, P.: Covid-ct-dataset: a ct scan dataset about covid-19. arXiv preprint arXiv:2003.13865 (2020)
113. Yao, Q., Xiao, L., Liu, P., Zhou, S.K.: Label-free segmentation of covid-19 lesions in lung ct. IEEE Transactions on Medical Imaging (2021)
114. Yao, Y., Liu, F., Zhou, Z., Wang, Y., Shen, W., Yuille, A., Lu, Y.: Unsupervised domain adaptation through shape modeling for medical image segmentation. arXiv preprint arXiv:2207.02529 (2022)
115. Yee, P.P., Li, W.: Tumor necrosis: A synergistic consequence of metabolic stress and inflammation. Bioessays **43**(7), 2100029 (2021)
116. Yoon, D., Kong, H.J., Kim, B.S., Cho, W.S., Lee, J.C., Cho, M., Lim, M.H., Yang, S.Y., Lim, S.H., Lee, J., et al.: Colonoscopic image synthesis with generative adversarial network for enhanced detection of sessile serrated lesions using convolutional neural network. Scientific reports **12**(1), 261 (2022)
117. Yu, B., Zhou, L., Wang, L., Frapp, J., Bourgeat, P.: 3d cgan based cross-modality mr image synthesis for brain tumor segmentation. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), pp. 626–630. IEEE (2018)
118. Zhang, Y., Li, X., Chen, H., Yuille, A.L., Liu, Y., Zhou, Z.: Continual learning for abdominal multi-organ and tumor segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 35–45. Springer (2023). URL <https://github.com/MrGiovanni/ContinualLearning>
119. Zhao, J., Li, D., Kassam, Z., Howey, J., Chong, J., Chen, B., Li, S.: Tripartite-gan: Synthesizing liver contrast-enhanced mri to improve tumor detection. Medical image analysis **63**, 101667 (2020)
120. Zhou, Z.: Towards annotation-efficient deep learning for computer-aided diagnosis. Ph.D. thesis, Arizona State University (2021). URL <https://github.com/MrGiovanni/Dissertation>
121. Zhou, Z., Gotway, M.B., Liang, J.: Interpreting medical images. In: Intelligent Systems in Medicine and Health, pp. 343–371. Springer (2022)
122. Zhou, Z., Sodha, V., Siddiquee, M.M.R., Feng, R., Tajbakhsh, N., Gotway, M.B., Liang, J.: Models genesis: Generic autodidactic models for 3d medical image analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 384–393. Springer (2019). URL <https://github.com/MrGiovanni/ModelsGenesis>
123. Zhu, Z., Kang, M., Yuille, A., Zhou, Z.: Assembling and exploiting large-scale existing labels of common thorax diseases for improved covid-19 classification using chest radiographs. In: Radiological Society of North America (RSNA) (2022). URL <https://github.com/MrGiovanni/LabelAssemble>