



Towards a better understanding of annotation tools for medical imaging: a survey

Manar Aljabri¹ · Manal AlAmir¹ · Manal AlGhamdi¹  · Mohamed Abdel-Mottaleb² · Fernando Collado-Mesa³

Received: 6 April 2021 / Revised: 4 August 2021 / Accepted: 3 January 2022 /

Published online: 25 March 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Medical imaging refers to several different technologies that are used to view the human body to diagnose, monitor, or treat medical conditions. It requires significant expertise to efficiently and correctly interpret the images generated by each of these technologies, which among others include radiography, ultrasound, and magnetic resonance imaging. Deep learning and machine learning techniques provide different solutions for medical image interpretation including those associated with detection and diagnosis. Despite the huge success of deep learning algorithms in image analysis, training algorithms to reach human-level performance in these tasks depends on the availability of large amounts of high-quality training data, including high-quality annotations to serve as ground-truth. Different annotation tools have been developed to assist with the annotation process. In this survey, we present the currently available annotation tools for medical imaging, including descriptions of graphical user interfaces (GUI) and supporting instruments. The main contribution of this study is to provide an intensive review of the popular annotation tools and show their successful usage in annotating medical imaging dataset to guide researchers in this area.

Keywords Annotations · Medical images · Segmentation · Tagging

1 Introduction

The relatively recent development of very powerful computational hardware, such as graphic processing units (GPUs), and the development of deep neural networks, paired with the availability of large quantities of digital data, have facilitated for machine learning

Manar Aljabri and Manal AlAmir contributed equally to the work.

✉ Manal AlGhamdi
maalghamdi@uqu.edu.sa

¹ Department of Computer Science, Umm Al-Qura University, Mecca, Saudi Arabia

² Department of Electrical and Computer Engineering, University of Miami, Florida, FL, USA

³ Department of Radiology, University of Miami Miller School of Medicine, Florida, FL, USA

(ML) to emerge as a field with the potential to generate great progress in different fields. ML is a subfield of artificial intelligence (AI) which includes a wide range of computational algorithms and modeling tools utilized to process large numbers of data, where these algorithms aim to mimic human intelligence by learning from training data. ML has been applied to different fields including robotics, pattern recognition, data mining, object recognition, face detection, and medical diagnosis [13, 48, 53, 85, 115].

Deep learning (DL) is a subset of ML, which aims at learning many levels of distributed representations of the data to be modeled. DL models have achieved massive progress in terms of algorithms, applications, and theories. It utilizes hierarchical recombination of features to extract relevant information and then learn the pattern representation by employing a multi-layer neural network. In recent years, medical image analysis has been boosted by ML and DL. The ML and DL methods help doctors diagnose the disease, predict its risk, and prevent them at an appropriate time [34]. It also helps to predict the number of patients of upcoming days in pandemics [86]. However, DL- and ML-based applications usually demand a large amount of annotated data to train the model, which relies on human experts with specialized knowledge and clinical experience to achieve adequate results. The model's performance improves as the amount of annotated data increases. These annotated data are known as Ground-Truth (GT) and used for training, testing, and evaluating the models. GT represents the optimal performance that an algorithm is desired to achieve [20]. As illustrated in Fig. 1, the ML process starts with the raw dataset which needs annotation for training and testing the models. For evaluating the algorithmic performance, the deviation of a predicted result is measured with respect to the appropriate GT.

Good annotation tools are in great demand due to the massive increase in digital data [31]. The annotation process aims to transfer human knowledge to the artificial intelligence models by summarizing and assigning predefined labels to the digital data content [9, 101]. Annotation tools are characterized based on the covered tasks, the provided functionalities, and the supported features as pre-processing and automatic labeling [87]. These tools enable the user to label an interesting object in a frame by supporting three modalities: manual, semi-automatic, and automatic [19, 31].

Manual annotation of audio, digital image, text, or video is the initial processing step of most research projects and systems [42]. It requires human annotators to delineate and label spatial regions in the image or define the temporal segments related to audio or video. The spatial regions are specified utilizing a standard shape, *i.e.*, circle, point, freehand-drawn mask, ellipse, polygon, polyline, *etc.*, whereas the temporal segments are determined by beginning and end timestamps. For example, a facial landmark helps in annotating the human face for detecting the position of the key point in the face such as nose tip, eye corners, and eyebrows [40]. These landmarks are the main components in different face applications including face recognition, facial attribute analysis, and face verification [33, 56]. Fig. 2 illustrates the most used annotation techniques for labeling images and text.

In most image-processing tasks, the desired annotations may range from labels at the image level (image classification), to framing boxes (object detection) or annotation at a

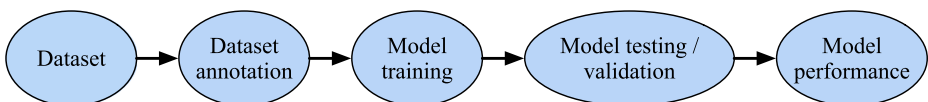


Fig. 1 Phases in the machine learning process. It starts with collecting raw data that demands annotation to train the model and measure its performance

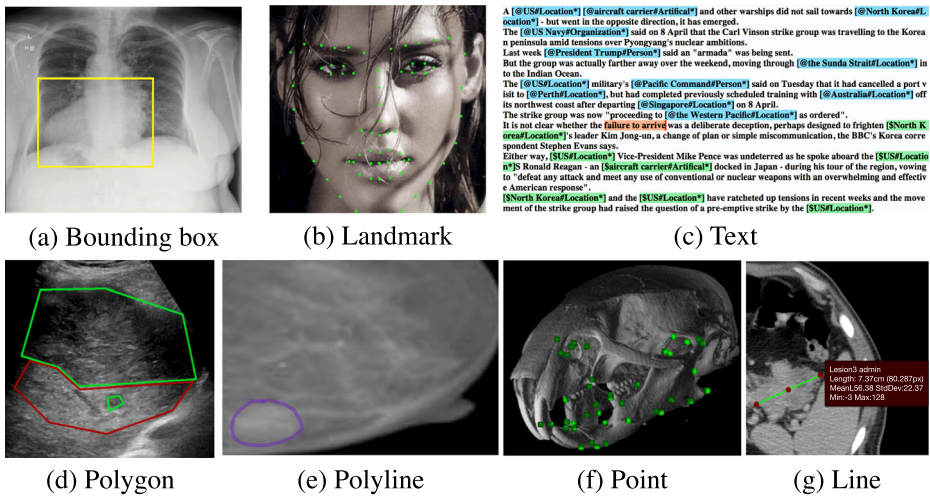


Fig. 2 Common annotation technique used in the most tools. (a) Disease bounding box in chest X-ray image from [92]. (b) Facial landmark location apply in grayscale image from [40]. (c) Text annotation by assigning the label to the part of the text to be helpful for Named Entity Recognition (NER) from [152]. (d) Polygon annotation in ultrasound image from [21]. (e) Polyline annotation by using Supervisely tool for X-ray image from [26]. (f) Point annotation of the 40 mandible landmarks [22]. (g) Line annotation of a cancer lesion [122]

pixel level (image segmentation) [38]. Image annotation and segmentation are core components for computer-aided diagnosis (CAD) and image recognition systems [102]. CAD systems use medical images to identify image features and diagnose lesions [116]. Moreover, CAD recognizes regions of interest (ROIs) by utilizing image segmentation and automatic annotation tools, to identify the relevant region. Medical imaging refers to several different technologies that are used to view the human body to diagnose, monitor, or treat medical conditions. It requires significant expertise to efficiently and correctly interpret the images generated by each of these technologies, which among others include radiography, ultrasound, and magnetic resonance imaging. Image annotation is widely employed in medical applications, where imaging modality is annotated by an expert to improve the model’s performance for lesion and disease detection. It is an approach to identify regions and add descriptions, explanations, or comments of these regions in a textual form. Annotating medical images should be accurate, therefore, usually, multiple experts view the data separately and perform the manual annotations [107]. After that, each annotator views the data independently and updates the annotations with primitive tagging. Next, they judge the annotations together and agree to any changes to correct and update the annotations. This process is known as the gold-standard annotation, see Fig. 3.

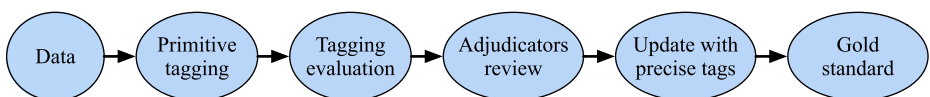


Fig. 3 Annotation process to achieve high accuracy annotation and satisfy the gold standard technique

Defining the appropriate annotation tools saves time and effort, however, knowing all existing tools and choosing the most suitable one among them is much complicated. Moreover, many annotation tools have emerged last few years for various tasks, this motivates us to provide a high-level glance of their successful usage, graphical user interface (GUI), available annotation techniques, and supported features, *i.e.*, zooming, input, and output. The main intention of this paper is to explore the major annotation tools for medical image tagging.

The remainder of the paper is structured as follows. A comprehensive review of related surveys is presented in Section 1.1. Section 2 presents the possible annotations for medical images including input formats and exportation. Section 3 presents the annotation tools applied for the medical images with its snapshot of these tools. Medical image applications that employed the reviewed tools are described in Section 4. Lastly, Section 5 discusses the reviewed tools and Section 6 concludes the paper.

1.1 Related survey

Previous surveys devoted to annotation tools offer a synopsis of the whole field while presenting the supported feature. In recent years, a significant amount of research has emerged about the use of different annotation tools for different content types. This illustrates the important role of annotation tools in the modern world. Dasiopoulou et al. [35] reviewed image and video annotation tools concerning functionality and interoperability perspectives, to focus on the problem of the communication, sharing, and reuse of generated metadata. Neves and Leser [105] presented a survey of biomedical text annotation tools, featuring 35 criteria to evaluate 13 annotation tools. These criteria encompassed issues of documentation, supported formats, extensibility, implemented functionality, platforms, and popularity. Gaur et al. [50] presented a survey for five current used annotation tools for video tagging with snapshots of tools' GUIs. These tools are VATIC, Beaverdam, ViTBAT, iVAT, and MViPER-GT. Moreover, they compared these tools in terms of platforms, targets, object shape, the used machine learning algorithms, and interface design. Rebinth and Kumar [116] reviewed various manual annotation tools and different available datasets. They presented the used tools in the segmentation stage for medical imaging to provide automatic detection and diagnosis of diseases.

In this survey, we focus on the tools that support manual annotation for medical imaging, which has been proven its successful application for generating a gold standard annotation by at least one research. Table 1 compares our survey to others in terms of the medical image tools, tools' snapshots, and mentioned applications of the tools. We selected the tools according to the following constraints: the tools should be easy to use, publicly available and support zooming. This survey will focus on the medical field and discuss the 13 most popular tools in tagging images. Moreover, we will support the tools with GUI snapshots and their successful application.

2 Medical image annotation

Image annotation is the process of classifying or labeling an image using text, annotation tools or both, to make a set of corresponding labels for each image to train the ML and DL models. This process is commonly applied to identify objects, boundaries and to segment images. Therefore, medical image annotation is the process of labeling medical images from different imaging modalities such as MRI, CT Scan, Ultrasound, Mammography *etc.*, for

Table 1 Comparison between this survey and the related ones in terms of number of reviewed tools, medical image applications, tools' snapshots and mentioned tools' application

Publication	Year	# tools	Medical	Snapshots	Applications	Remarks
Dasiopoulou et al. [35]	2014	14	✗	✓	✗	image and video annotation tools concerning functionality and interoperability perspectives.
Neves and Leser [105]	2012	13	✗	✗	✓	A synopsis of biomedical text annotation tools.
Gaur et al. [50]	2018	5	✗	✓	✗	A summary of five annotation tools for video.
Kumar [116]	2019	22	✓	✗	✗	Various manual annotation tools with medical datasets.
This survey	2021	13	✓	✓	✓	Presenting the successful usage of annotation tools in medical domain with applications.

ML and DL training. These annotation play a significant role in the healthcare sector to assist with diagnosing different diseases, segmenting organs at risk before radiation therapy and performing robotic surgery.

Figure 4 shows different types of medical image annotation. The segmentation task for different human body organs allow further quantitative analysis of many clinical parameters, including shape and volume such as in cardiac or brain image analysis. In addition, it is often a significant first step in CAD pipelines as shown in Fig. 5.

Many ML and DL technics in the medical field aim to segment or detect abnormalities to quantify or classify them into malignant or benign. Both segmentation and classification processes can be considered as a classification task, first classifying each pixel as belonging to a lesion or not, which is known as a semantic segmentation. Then, classifying segmented abnormalities as malignant or benign [77]. The detection of objects or lesions in medical images is an important part of disease diagnosis, but it is often a labor-intensive process. In most cases, the detection consists of the localization and identification of a small part in the full image. There has been extensive research on CAD systems that are developed to automatically detect lesions with high accuracy and thus decrease the reading time of human experts. Annotation tools have been used to segment different objects in different organs, such as masses in the breast, nodules in the lungs, vessels in the retina, and tumors in the liver, brain and other organs as shown in Fig. 6.

2.1 Input and output

Various scanning techniques have been used to visualize the interior of the human body generating multiple modalities including X-ray such as in mammography [3], Ultrasound (US), Magnetic Resonance imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), microscopy images, histology slide images, dermoscopy images, Optical Coherence Tomography (OCT) images, and color fundus images [8]. Figure 7 shows examples of medical images. CT and MRI are able to examine multiple organs at the same time, while retinal photography and dermoscopy are organ-specific. The amount of data generated from each imaging modality vary, for instance an MRI can be a few hundred megabytes while a histology slide is an image file of a few megabytes. This has technical effects on the way the data is pre-processed, and on the design of the model architecture,

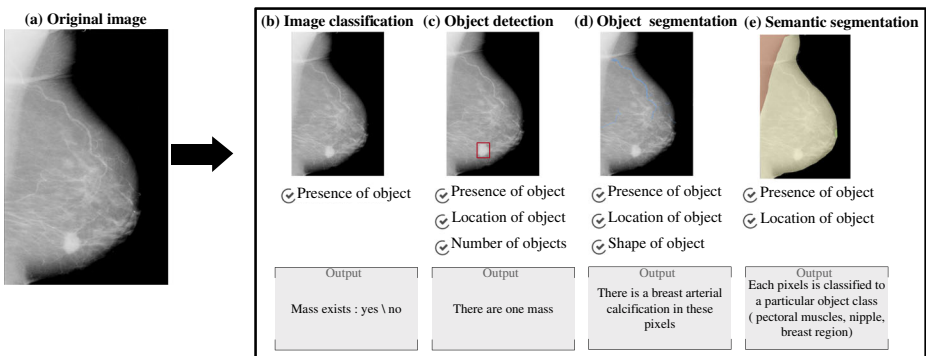


Fig. 4 Different medical image annotation types (a) Original image [61]. (b) Benign and malignant classification of mammogram images. (c) Mass detection. (d) Breast arterial calcification segmentation in mammogram images. (e) Segment of different breast tissue regions in mammograms: Four regions (pectoral muscle; fatty region; nipple region; dense region)

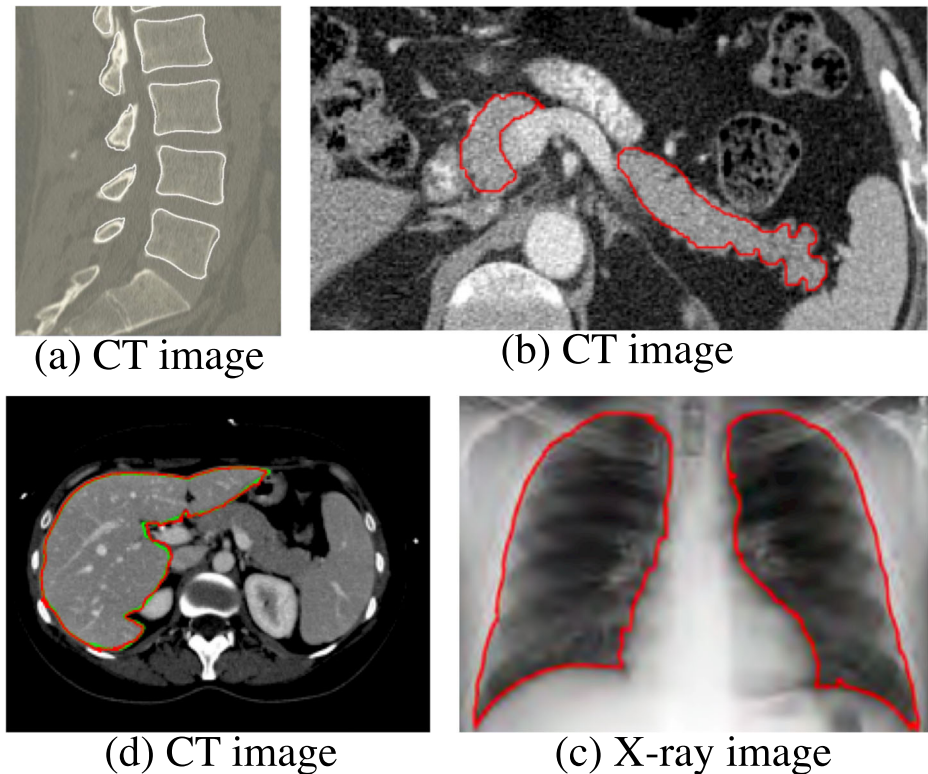


Fig. 5 Organ and substructures annotation: (a) Vertebrae segmentation [114]. (b) Pancreas segmentation [119]. (c) Lung segmentation [23]. (d) Liver segmentation [41]

in terms of processor and memory limitations. Medical imaging has led to improvements in the diagnosis and treatment of numerous medical conditions in children and adults.

There are several of medical imaging, each of which uses different technologies and techniques. CT and radiography (including mammography) use ionizing radiation to generate images of the body. In radiography a single image is recorded for later evaluation (mammography is a special type of radiography to image the internal structures of breasts), while in CT many X-ray images are recorded as the detector moves around the patient's body, a computer reconstructs all the individual images into cross-sectional images or 'slices' of internal organs and tissues. 2D images generated by X-ray are used in several evaluation settings such as bone fractures, pneumonia, pulmonary edema, renal or gallbladder stones and intestinal obstructions. CT images are used in many settings, including trauma evaluation, and can be used to evaluate internal organ systems such as neurologic, gastrointestinal, genitourinary, and the vascular system. MRI is a medical imaging procedure for making images of the internal structures of the body. MRI scanners use strong magnetic fields and radio waves (radiofrequency energy) to make images. During an MRI exam, an electric current is passed through coiled wires to create a temporary magnetic field in a patient's body. Radio waves are sent from and received by a transmitter/receiver in the machine, and these

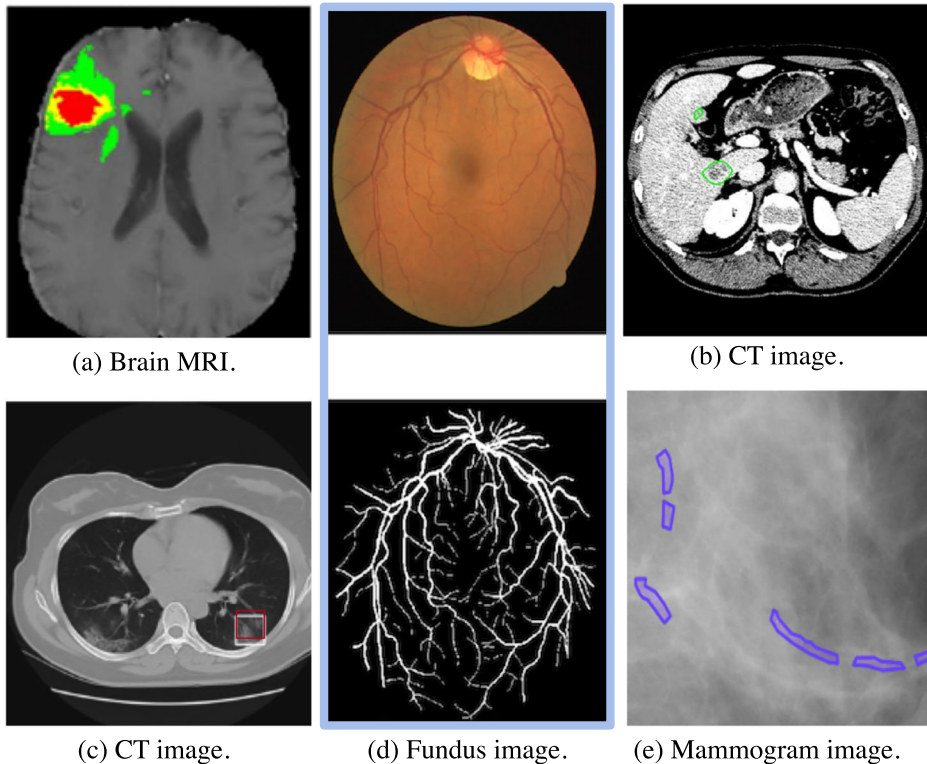


Fig. 6 Part of organ annotation: (a) Brain tumour segmentation [110]. (b) Liver tumour segmentation [135]. (c) Lung lesions detection [151]. (d) Retinal vessels segmentation [134]. (e) Breast arterial calcification segmentation [5]

signals are used to make digital images of the scanned area of the body. OCT images are the main diagnostic technology for retinal diseases. Fundus images show the retina, the optic disk and blood vessels. It is used in to diagnose diabetes retinopathy, macular degeneration, and glaucoma.

Medical image file formats can be divided into two main groups. The first format is intended to standardize the images generated by diagnostic modalities like Dicom. The second format is aimed to facilitate post-processing analysis like Analyze, Nifti, and Minc [88]. Some of the annotation tools support these formats while others require to convert the file to an image file format such as PNG or JPEG.

Annotation outputs are usually exported in one of the following file formats: Comma Separated Value (CSV) files, text files, Java Script Object Notation (JSON), Tensorflow Records (TFRecord), or as database-specific files [143]. JSON is a lightweight data-interchange format, which is considered easy for humans to read and write and also for machines to parse and process. Fig. 8 shows an example of expected output files for mass detection in mammography images. TFRecord file stores data as a sequence of binary records. Commonly used database-specific files include COCO which stores annotations using JSON, Pascal VOC that stores annotations in XML file, and YOLO that stores annotation in .txt file.

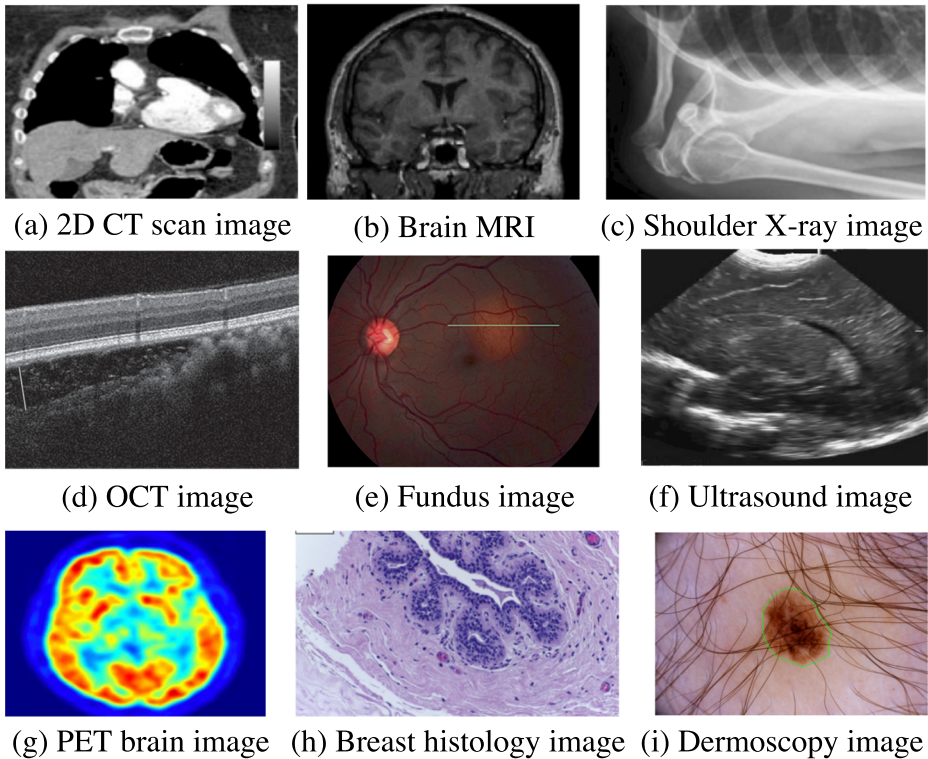


Fig. 7 Examples of expected inputs figures are cropped from the corresponding papers: (a) 2D CT scan image [76], (b) Brain MRI image [76], (c) Shoulder X-ray image [76], (d) OCT image [11], (e) Fundus image [11], (f) Ultrasound image [76], (g) PET brain image [76], (h) Breast histology image [10], (i) Dermoscopy image [153]

3 Annotation tools

Choosing the most appropriate tool for a specific application is very important because it significantly affects the quality of the data and the time needed to complete it. For this section we researched and tested 13 different image annotation tools and summarized the features of each one. In addition, we make reference to the papers where these tools were used to accomplish different tasks. For all the reviewed tools, we show a snapshot of the tools' GUI where we tested the tool by performing a segmentation on a mammogram image to annotate different objects, including the pectoral muscle, fatty tissue, the nipple, and a breast mass.

3.1 VGG Image Annotator (VIA)

VIA¹ [42] is a software developed using HTML, CSS, and JavaScript by the Visual Geometry Group (VGG) from the University of Oxford. VIA is open source, straightforward,

¹<https://annotate.officialstatistics.org/>

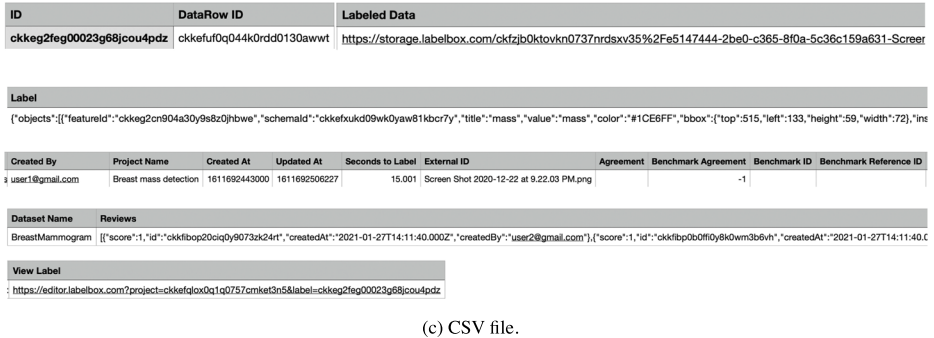


Fig. 8 Examples of expected outputs: (a) is the corresponding JSON file for (b) the detected mass in breast mammogram, (c) is the corresponding CSV file for (b) the same detected mass

available for free and a manual annotation software used for video, image, and audio segmentation. It does not require any installation or setup since it runs in a web browser. This software adapts into a single self-contained HTML page with a size less than 400 Kilobyte, which operates as an offline program in most modern web browsers. The annotations export as JSON and CSV to allow further processing by other software tools. The software provides cooperative annotation of a big dataset by a group of human annotators. Also, it does not dependent on the external libraries. For detecting regions, it offers six shapes: polygon, rectangle, ellipse, circle, polyline, and point. The most used one is the rectangle, which is suitable for determining the object bounding box. The point is used to specify feature points like a landmark or the main points on MRI images. The text description is used to demonstrate the region content. A snapshot of the VIA’s GUI is presented in Fig. 9. This tool has been used for semantic segmentation [6], video annotation [2, 129], and image annotation [12, 25].

3.2 LabelIMG

LabelIMG² is a software developed using Python and QT utilized for the graphical interface. It is an open-source, straight-forward, offline tool that can be accessed from both Windows

²<https://github.com/tzualin/labelimg>

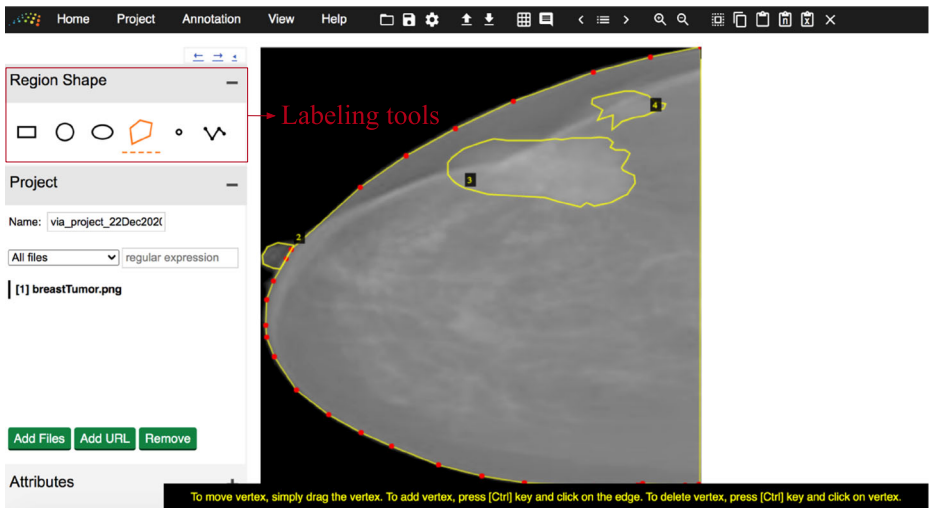


Fig. 9 Breast parts and tumor annotation using VIA tool in X-ray image [26]

and Mac operating systems via the GitHub. It annotates objects in graphical images using only bounding boxes. Annotations are stored as XML files in PASCAL VOC format, and it also supports YOLO format. However, it only supports a straightforward image annotation function which means that it does not provide annotation for a stream of the image, and without auxiliary annotation [154]. Figure 10 shows the snapshot of LabelIMG's GUI. This tool was previously used for annotating ship images [18] and labeling bird images [118].

3.3 Ratsnake

Ratsnake³ [67] is a generic image annotation, semantically aware software developed using Java. It is open source, straightforward, and provides quick image annotation with snakes. Ratsnake uses a semi-automatic approach for graphical image annotation and depends on a customizable active contour model with consideration of rapid user annotations. It allows fast segmentation and annotation of images with polygons, grids, or both. Moreover, it transforms binary masks to polygon annotations. However, only one mask can be added for each image. Ratsnake supports exporting an annotation as custom text, LabelMe XML, and owl format. Furthermore, it requires Java Virtual Machine to be installed. Figure 11 shows the snapshot of Ratsnake's GUI. This tool was previously used for annotating images obtained by the wireless capsule [70] and annotating coherent regions in video [139].

3.4 Visual object tagging tool (VOTT)

VOTT⁴ is a software developed using TypeScript by the Commercial Software Engineering (CSE) group at Microsoft. It is an open-source, straightforward web application, available for free on GitHub and used for graphical images and video. It annotates objects in image

³<https://is-innovation.eu/ratsnake/>

⁴<https://github.com/microsoft/VoTT>

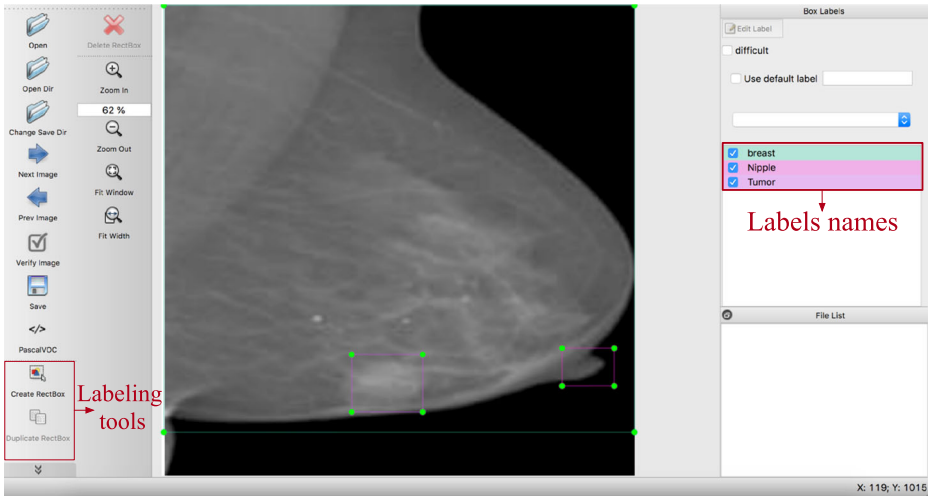


Fig. 10 Breast parts and tumor annotation using LabelIMG tool

using only bounding boxes. Annotations are exported in many formats as Microsoft Cognitive Toolkit TensorFlow (Pascal VOC and TFRecords), VoTT (generic JSON schema), and CSV. It provides the ability to export and import from local or cloud storage. Fig. 12 shows the snapshot of VOTT's GUI. It was previously used for annotating feather images [15] and for annotating images in the YawDD dataset [1] in [4].

3.5 Mask editor

Mask Editor⁵ [160] is a software developed using MATLAB. It is an open source, straightforward, and available free for use to generate image mask. It supports drawing irregular mask shapes around the object. It provides many annotation functions including erasing, super-pixel-marking, cropping, zooming, navigation between images and B-curve. It saves images with multiple formats including JPG, BMP, TIFF, and PNG. Figure 13 shows the snapshot of Mask Editor's GUI. Mask Editor was previously used for annotating surgical tools [31].

3.6 Supervisely

Supervisely⁶ [136] was developed by Deep Systems as a powerful platform for computer vision development, where individual annotators and large teams can work together and experiment with datasets. It is a web-based application with an easy-to-use GUI that helps individuals with and without ML experience to create computer vision applications. Supervisely has tools to draw the one in a completely manual way or semi-automatic way by selecting the desired area to create the marking and automatically generating the desired shape. It has quick access commands to make the marking process more efficient. Another

⁵<https://github.com/Chuanhai/Mask-Editor>

⁶<https://supervisely/>

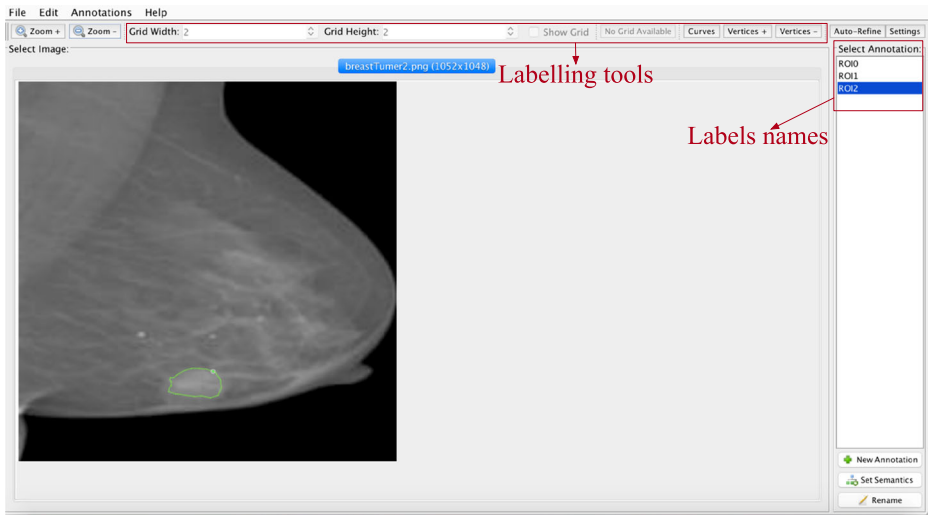


Fig. 11 Breast tumor annotation using Ratsnake tool

important function is the ability to modify contrast and brightness to improve the marking process. Also, it can annotate either with vector graphics or on pixel level. The vector graphics tools are polygon, rectangle, polyline, and point, while the pixel level tools are brush, eraser, and smart tool. It enables the user to perform different functions regarding geometric objects, labeling data, and tags. It supports images, videos, volumetrics, and medical data with various formats including .png, .jpeg, .mp4, .avi, .dicom, .pcd and

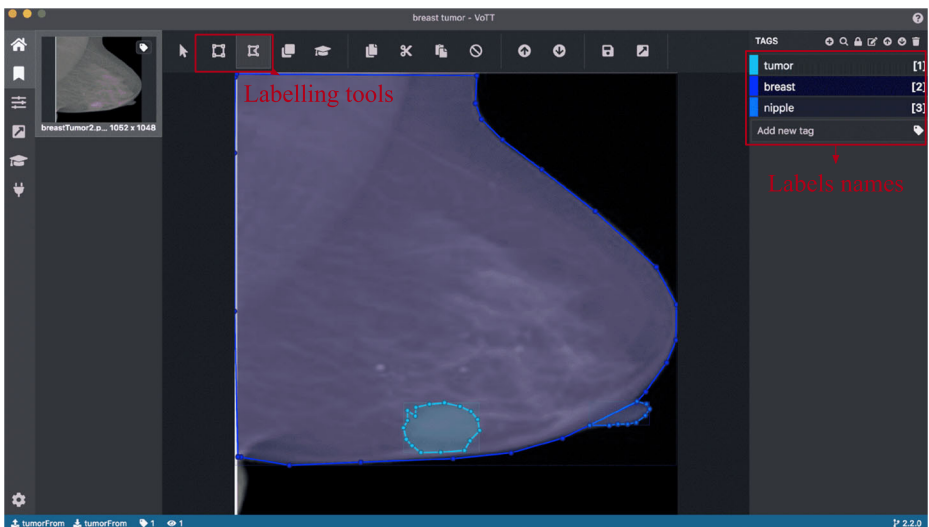


Fig. 12 Breast parts and tumor annotation using VOTT tool in X-ray image

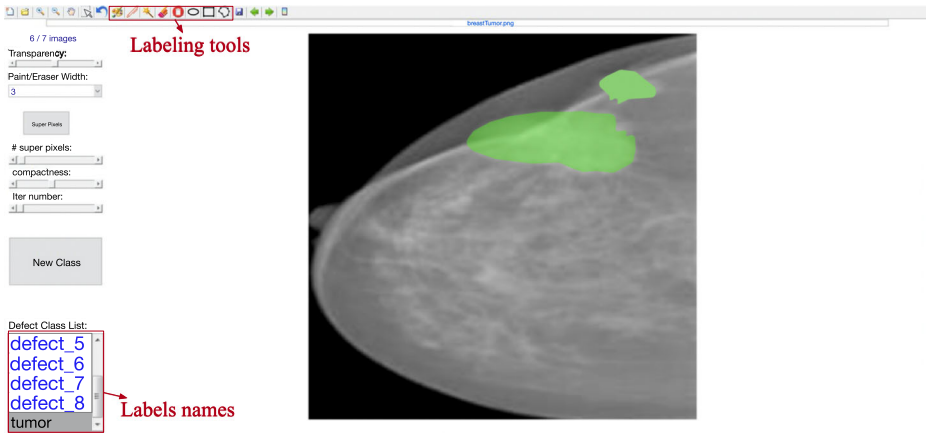


Fig. 13 Breast tumor annotation using Mask Editor tool

others. Furthermore, there are multiple ways for the annotations to be export like .json, .png masks, .trecords, .xml and more. However, time statistics and quality control mechanisms are missing in Supervisely. This community edition of this tools is free but a price is charged for self-hosted versions. Fig. 14 shows a snapshot using the Supervisely. This tool has been used before for object segmentation [14, 63], bone segmentation [71], object detection [46, 54, 108, 111, 117], segmentation [90, 158, 159] and segmentation of pneumothorax [138].

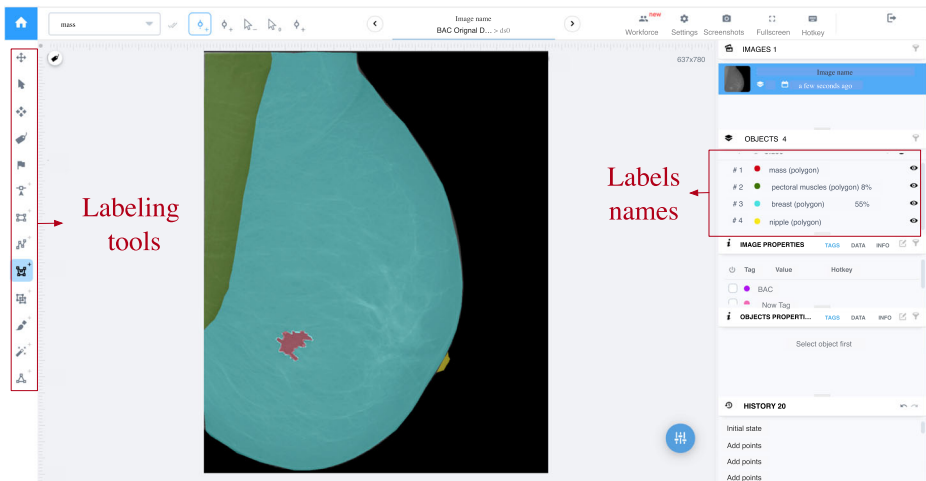


Fig. 14 Breast parts and tumor annotation using the Supervisely tool



Fig. 15 Breast parts and tumor annotation using RectLabel tool

3.7 RectLabel

RectLabel⁷ [78] (2017) is an image annotation tool for labelling images used for object detection, bounding box, and segmentation. There are many features of RectLabel like drawing a polygon, bounding box, line, and cubic bezier. Rectlabel allows drawing key points and the skeleton, label image pixels with the brush. It provides automatic superpixel tools to label images. RectLabel allows reading and writing in PASCAL, VOC and XML formats and enables the user to export to YOLO, COCO JSON, and CSV formats. It also allows users to export index color mask and separated mask images. RectLabel provides user- friendly labeling of images and retrieves images based on labels. However, it is only available through the Mac App Store. Fig. 15 shows a snapshot using the RectLabel. This tool has been used before for object detection [73, 79, 147, 148], Vessel lumen segmentation [150] and detection Polyps [112].

3.8 LabelMe

LabelMe⁸ [123] is a free, web-based software developed by MIT Computer Science and Artificial Intelligence Laboratory using JavaScript. This tool enables users to annotate images, focusing on ease of use and simplicity of the design. The segmentation is performed by drawing polygons over the objects of interest on the images. The user can export results in an XML file format, making them easy to extend and transfer. Fig. 16 shows a snapshot of LabelMe GUI. This tool has been used before for object segmentation [58].

⁷<https://rectlabel.com/>

⁸<http://labelme2.csail.mit.edu/Release3.0/index.php>

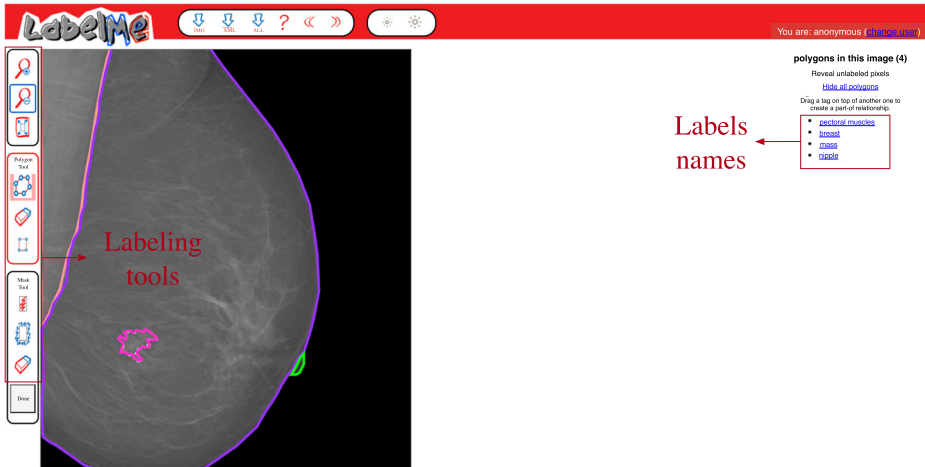


Fig. 16 Breast parts and tumor annotation using LabelMe tool

3.9 Labelme

Labelme⁹ [144] was developed by Kentaro Wada based on the previous Labelme (Section 3.8). It is a graphical image and video annotation tool, which is written in Python and uses Qt for its graphical interface. It provides polygon, rectangle, circle, line and point tools. It can be downloaded in Ubuntu, MacOS and Windows. Labelme enables the user to select a region to annotate the rectangular box which contains the object. After this, they can add a category or label of the object contained in the box. Finally, the annotation files can be exported as JSON files in VOC or COCO formats. The polygon annotation tool has a detailed contour of the object, which can be useful for image segmentation task. The disadvantage of LabelMe is that images can be only accepted in JPG format. Fig. 17 shows a snapshot of Labelme GUI. This tool has been used before for object segmentation [28, 47, 96, 121, 145] and object detection [80].

3.10 Computer vision annotation tool (CVAT)

CVAT¹⁰ [72] is an open-source program developed by Intel to annotate both images and videos data. The process in CVAT starts by creating an annotation task with a specific name, labels, and attributes. Datasets are then loaded from a mounted file system inside a container or the local file system. A task can include one image, one video or a set of images from shared storage. It allows users to annotate images with several types of shapes like boxes, polygons which used for both general and segmentation tasks, polylines, and points. CVAT is easy to reach by a web-based interface. However, CVAT has only been tested in Google Chrome browser and may not work well in other browsers. CVAT supports different image

⁹<https://github.com/wkentaro/labelme>

¹⁰<https://cvat.org>



Fig. 17 Breast parts and tumor annotation using Labelme tool

and video formats like *.png, *.jpg, *.mp4 and enables users to export annotations and images in a specific format such as CVAT for video, CVAT for images, PASCAL VOC and many other dataset formats. Fig. 18 shows a snapshot of CVAT GUI. This tool has been used before for object detection [106, 155] and segmentation [132].



Fig. 18 Breast parts and tumor annotation using CVAT tool

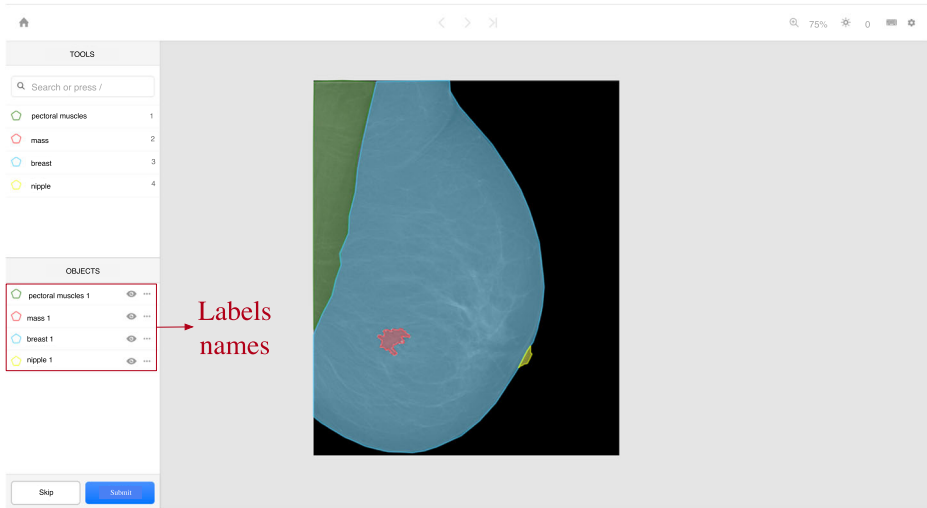


Fig. 19 Breast parts and tumor annotation using LabelBox tool

3.11 LabelBox

LabelBox¹¹ developed by Sharma, Daniel Rasmuson and Brian Rieger [127] is a free commercial online web-based annotation system for segmentation and classification purposes. It includes different types of marker, which are a line, point, brush, superpixel and brush. After finishing the annotation task, we can export the mask results in different formats such as CSV and JSON. The generated mask is compatible with multiple models such as Mask RCNN. LabelBox has one of the best user experiences so far. One of the things that makes annotation easier on LabelBox is that when we draw a marker on the object in an image, the polygon will move to the object border. Fig. 19 shows a snapshot of Labelbox. This tool has been used before for object detection [99] and object segmentation [5, 30, 62].

3.12 ITK-SNAP

ITK-SNAP¹² [156] is a software application that allows users to annotate 3D medical images, manually draw anatomical areas, and automatically perform image segmentation. It was designed with the audience of the science and clinical researchers in mind; thus a focus has been put on providing a user-friendly interface and keeping a limited feature to prevent feature creep. ITK-SNAP is mostly used to work with Cone-Beam Computed Tomography (CBCT), MRI and CT data. The main features of the software are manual segmentation, image navigation, and automatic segmentation. ITK-SNAP is an open-source, free and multi-platform. It supports many different 3D image formats, like NIFTI and DICOM and export the segmentation results as images. Fig. 20 shows a snapshot of ITK-SNAP. This tool has been used before for 3D object segmentation [17, 157].

¹¹<https://labelbox.com>

¹²<http://www.itksnap.org/pmwiki/pmwiki.php>

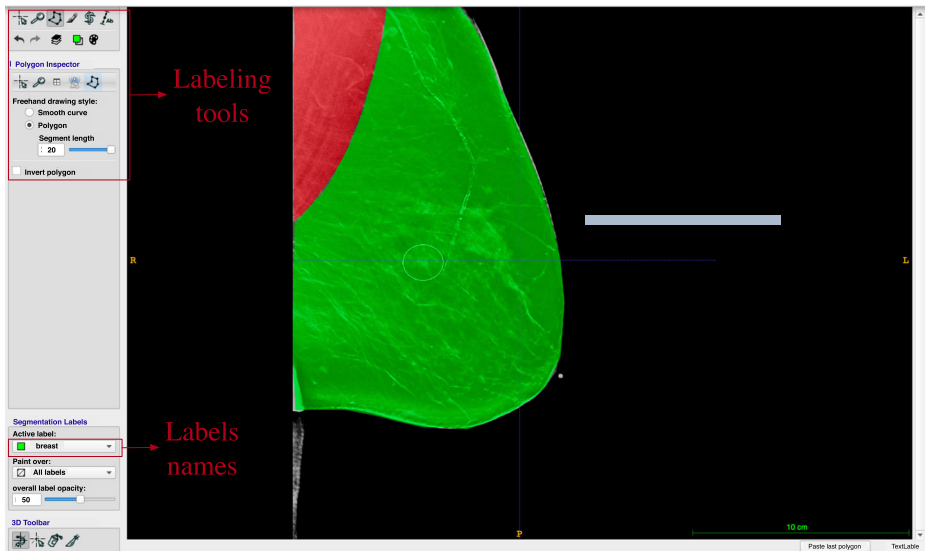


Fig. 20 Breast parts annotation using ITK-SNAP tool

3.13 3D-Slicer

3D-Slicer¹³ [44] is a multi-module software in which each module performs a special 3D medical image processing task. It is a free and open-source software developed using Python and C++. For annotating and segmentation of medical 3D images there are many modules that can be helpful for instance including Simple Region Growing Segmentation that is based on intensity statistics, EMSegment Easy that performs a quick intensity-based image segmentation on MRI and the Editor module that includes a collection of tools for manual segmentation (*e.g.*, paint, draw) and semi-automatic segmentation (*e.g.*, thresholding, region growing, interpolation). 3D-Slicer allows users to upload an extensive variety of image formats and includes converting format functions. Output of segmentation can be exported as NRRD or NIFTI file. Fig. 21 depicts a snapshot of 3D-Slicer. This tool has been used before for 3D object segmentation [124, 128, 141, 161].

4 Medical application

For developing a robust and precise model, medical image annotation is required, but also considered a major hurdle [120]. Image-based CAD models try to facilitate medical lesion detection and abnormalities by evaluating medical images as objectively as possible, using image features and prior knowledge about the particular application domain. Such systems usually combine image segmentation methods to isolate ROIs corresponding to prominent objects, and automatic annotation methods, to attach labels that characterize each region. Prior knowledge is typically acquired from associated medical studies and domain experts, by manual annotation and segmentation of images [67]. In this section, we will present some

¹³<https://www.slicer.org/>

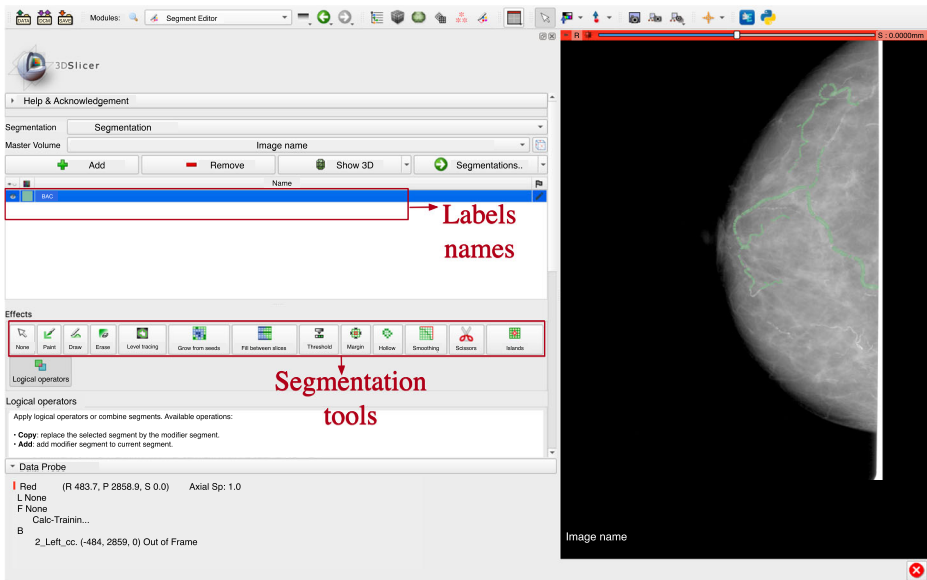


Fig. 21 Breast arterial calcification segmentation using 3D-Slicer tool

medical applications that employed the above-mentioned tools. Tools for both detection and segmentation tasks are summarized in Table 2.

4.1 Detection

Detection of abnormalities requires a tremendous amount of annotated data. Typically, detection is performed using a bounding box around each object of interest in an image [131]. The VIA tool has been employed for various detection tasks in medical images including knee joints detection by Kondal et al. [82], nuclei identification in placenta imaging by Ferlaino et al. [45], Chest region and heart detection in X-ray images by Gupteet al. [57] and foot ulcer detection by Cassidy et al. [24]. Additionally, it was utilized by Mallisery et al. [97] to identify sensitive data in medical images and by Rajaraman et al. [113] to detect abnormalities in chest radiographs related to COVID-19.

Various detection tasks have been developed based on LabelImg tool. For instance, osteoarthritis disease was detected in MRI images by Singh et al. [130], foot ulcers were detected by Cassidy et al. [24], other lesions were detected by Sha et al. [126] and malaria was detected by Nakasi et al. [104]. Additionally, Li et al. [91] used it for labeling tongue images. While Hahn et al. [60] used it for bounding abdominal aortic aneurysm region.

The Ratsnake tool was utilized for blood detection [68], and lesion detection [36, 69, 70]. VOTT was used for diagnosing brain tumors [43] and surgical tools detection [31]. Rahim et al. [112] used the RectLabel tool for detecting polyps in colonoscopy images. Wei et al. [146] used it to detect colorectal polyps on histopathology slides by setting rectangular bounding boxes around polyps. Moreover, Kawazoe et al. [79] employed it to detect glomerular in multi-stained human whole-slide histopathology images. Hadush et al. [59] used LabelMe to detect mass abnormalities in mammograms.

Table 2 Medical image segmentation summary. The table categorized according to the type of the image used in the research

Publication	Task	Used tool	Remarks
CT			
Gou et al. [55]	Segmentation	LabelMe	Teeth from CT image.
Liu et al. [94]	Segmentation	Labelme	Throat.
Chen et al. [27]	Segmentation	Labelme	Lungs.
SONG et al. [133]	Segmentation	Labelme	Gallstones.
Park et al. [109]	Segmentation	ITK-SNAP	Aneurysms.
Sanchez et al. [125]	Segmentation	ITK-SNAP	Bones.
Hahn et al. [60]	Detection	LabelImg	Using computerized tomography angiography image.
Sha et al. [126]	Detection	LabelImg	Spine.
Rusu et al. [124]	Segmentation	3D-Slicer	Lung.
Hu et al. [65]	Segmentation	3D-Slicer	Brain.
MRI			
Huang et al. [66]	Segmentation	LabelMe	Vertebral.
Yushkevich et al. [157]	Segmentation	ITK-SNAP	Brain.
Singh et al. [130]	Detection	LabelImg	Osteoarthritis disease.
Gaonkar et al. [49]	Segmentation	ITK-SNAP	Intervertebral disks.
Kim et al. [81]	Segmentation	ITK-SNAP	Supraspinatus fossa and muscle.
Mansoor et al. [98]	Segmentation	ITK-SNAP	Anterior visual pathway.
Ezhilarasi and Varalakshmi [43]	Detection	VOTT	Brain.
Vickery et al. [141]	Segmentation	3D-Slicer	Brain.
Choi et al. [29]	Segmentation	3D-Slicer	Brain.
Shen et al. [128]	Segmentation	3D-Slicer	Abdominal organs and adipose tissue compartments.
Muller et al. [103]	Segmentation	ITK-SNAP	Lower legs and thighs.

Table 2 (continued)

Publication	Task	Used tool	Remarks
X-ray			
Iglovikov et al. [71]	Segmentation	Supervisely	Hand bones.
Tofikachev et al. [138]	Segmentation	Supervisely	Pneumothorax air pockets in chest X-ray.
Vlontzos et al. [142]	Segmentation	LabelMe	Vessel and catheter from fluoroscopy X-Ray.
Kordon et al. [83]	Segmentation	Labelme	Femoral condyles from knee joint X-ray.
Tang et al. [137]	Segmentation	Labelme	Opacity regions in lung.
Gupte et al. [57]	Detection	VIA	Chest.
Xian et al. [149]	Segmentation	ITK-SNAP	Vessels from X-ray angiography.
Sirazitdinov et al. [132]	Segmentation	CVAT	Tubes, wires and catheters from chest X-ray.
Ultrasound			
Xie et al. [150]	Segmentation	Rectlabel	Vessel lumen.
Roy et al. [121]	Segmentation	Labelme	COVID-19 markers in lung.
Rajaraman et al. [113]	Detection	VIA	Lung.
Liu et al. [93]	Segmentation	ITK-SNAP	Liver tumor.
Brehar et al. [21]	Segmentation	VIA	Hepatocellular Carcinoma.
Vats et al. [140]	Segmentation	Ratsnake	Wireless capsule endoscopy images.
Mammography			
Hadush et al. [59]	Detection	LabelMe	Breast mass.
AlGhamdi et al. [5]	Segmentation	Labelbox	Breast arterial calcification.
Zhang et al. [161]	Segmentation	3D-Slicer	Breast (micro-calcification).
Retinal fundus imaging			
Yu et al. [154]	Segmentation	Labelme	Optic disc and the macula area.

Table 2 (continued)

Publication	Task	Used tool	Remarks
Histology image			
Wei et al. [146]	Detection	RectLabel	Colorectal polyps.
Kawazoe et al. [79]	Detection	RectLabel	Glomerular.
Other			
Kondal et al. [82]	Detection	VIA	Using knee radiograph for severity of knee osteoarthritis and convolutional neural network for object detection.
Ciarrone et al. [31]	Segmentation	VIA and Mask editor	Utilizing video frames for detecting surgical tools.
Rahim et al. [112]	Detection	RectLabel	Polyps from colonoscopy images.
Francois et al. [46]	Segmentation	Supervisely	Uterus from laparoscopy images.
Zadeh et al. [158]	Segmentation	Supervisely	Uterus, ovaries and surgical tools from laparoscopic gynaecological images.
Zaki et al. [159]	Segmentation	Supervisely	Nuclear from cell image.
Gurari et al. [58]	Segmentation	LabelMe	Lu melanoma cells from fluorescence imaging and rabbit aorta from MRI.
Zhu et al. [162]	Segmentation	LabelMe	Teeth from natural color images.
Gentil et al. [51]	Segmentation	LabelMe	Cells from microscopy image.
Ferlino et al. [45]	Detection	VIA	Placenta.
Besson et al. [17]	Segmentation	ITK-SNAP	Lung tumours from PET images.
Cassidy et al. [24]	Detection	VIA and Labelling	Foot.
Jha et al. [74]	Segmentation	Labelbox	Polyp from gastrointestinal polyp images.
Li et al. [91]	detection	Labelling	Tongue.
Nakasi et al. [104]	Detection	Labelling	Microscopic malaria diagnosis.
Iakovidis et al. [68]	Detection	Ratsnake	Using wireless capsule endoscopy images.
Alia et al. [7]	Segmentation	VIA	Video endoscopy.
Dhieb et al. [37]	Segmentation	VIA	Blood smear images.
Lee et al. [89]	Segmentation	VIA	Using 500 frames.
Koulaouzidis et al. [84]	Segmentation	Ratsnake	Capsule endoscopy images.
Deeba et al. [36]	Segmentation	Ratsnake	Endoscopic images.

4.2 Segmentation

Segmentation involves annotating the object at pixel-level detail. Ciaparrone et al. [31] utilized VIA for annotating surgical tools. Alia et al. [7] provided accurate semantic segmentation of endoscopy artifacts by using VIA. Dhieb et al. [37] developed a framework for automated blood cells using VIA to segment cell images. Additionally, Hosseini et al. [64] developed an approach for counting, detecting, and categorizing cells inside microscopy images and annotating images using VIA. Lee et al. [89] utilized it for segmenting surgical tools. Brehar et al. [21] annotated Hepatocellular Carcinoma (HCC), and cirrhotic parenchyma (PAR) in ultrasound images of the liver using VIA. Vats et al. [140] used the Ratsnake tool to identify ROIs in gastrointestinal images. It was also used to segment lesions [36, 84].

Iglovikov et al. [71] used the Supervisely tool for segmentation of hand bones to train a DL model to assess pediatric bone age. Moreover, Francois et al. [46] used it for semantic segmentation of laparoscopic images of the uterus to automatically detect the organ, including its contours. In addition, Tolkachev et al. [138] employed it to segment pneumothorax air pockets on X-ray. Supervisely was also used by Zadeh et al. [158] to segment ovaries, uteruses, and surgical tools from laparoscopic gynaecological images that were used in image-guided surgery systems. Zaki et al. [159] used it for manual segmentation of the nucleus from different cell types.

Xie et al. [150] used the RectLabel tool to segment the vessel lumen from ultrasound images. Gurari et al. [58] used LabelMe to segment several medical image datasets to evaluate expert, non-expert, and algorithm segmentation performances. Zhu et al. [162] used it to segment teeth from natural color images by creating polygons around the teeth. Gou et al. [55] used Labelme to segment teeth on CT images. While Vlontzos et al. [142] employed it to segment vessel and catheter from fluoroscopy images. Gentil et al. [51] used LabelMe to segment cells in microscopy images. It was also used by Huang et al. [66] to segment vertebrae from MR images.

Roy et al. [121] used the Labelme tool to segment COVID-19 markers in lung ultrasounds. Liu et al. [94] used it to segment the throat from CT images to consider the effect of the throat area on tracheal intubation difficulty according to the irregularity. Chen et al. [27] employed Labelme to segment lungs from CT image. While Kordon et al. [83] used it to segment femoral condyles from knee joint X-ray images. Yu et al. [154] segmented the optic disc and the macula area using Labelme tool. SONG et al. [133] segmented gallstones in CT images by physicians and experienced radiologists using Labelme. Tang et al. [137] used it to segment the opacity regions in the lung from X-ray for Covid-19 patients.

Yushkevich et al. [157] used the ITK-SNAP tool to segment multi-modality imaging datasets like MRI brain scans. Besson et al. [17] used it to segment lung tumours from fluorodeoxyglucose (FDG)- PET images. Gaonkar et al. [49] used ITK-SNAP to manually segment of intervertebral disks. Xian et al. [149] manually segmented the main vessels from X-ray angiography images by ITK-SNAP. Muller et al. [103] segmented muscle volumes of the left and right lower legs and thighs using ITK-SNAP software. Park et al. [109] using the ITK-SNAP to manually segment aneurysms on each slice of CT image to a diagnosis of cerebral aneurysms. Kim et al. [81] used it to segment two regions of supraspinatus fossa and muscle in MRI slices. Segmenting a liver tumor along its boundary from the ultrasound image using ITK-SNAP has been done by Liu et al. [93]. Mansoor et al. [98] used ITK-SNAP to segment an anterior visual pathway from MRI images. Sanchez et al. [125] used it to segment bones from CT images. AlGhamdi et al. [5] used Labelbox for breast

arterial calcification segmentation from mammography. Jha et al. [74] segmented polyps from gastrointestinal polyp images using Labelbox. Sirazitdinov et al. [132] used CVAT to manually annotate pixels which belong to tubes, wires, or catheters from a chest X-ray.

Vickery et al. [141] performed the annotation for the brain region using a 3D-Slicer. Besides, the 3D-Slicer was used to segment the hippocampal region [29], annotating abdominal organs [128], performing the data pre-processing for extracting the brain region [65], and highlighting a nodule in the chest [124]. Furthermore, Zhang et al. [161] employed a 3D-Slicer for positioning of the micro-calcification identified by a 3D bounding box in each digital breast tomosynthesis volume.

5 Discussion

The need to create annotated datasets has increased with the popularity and availability of application-specific DL methods. In the fact, building an extensive list of tools is challenging due to the large number of new tools that are being released while some of the old ones are no longer use. Moreover, defining the best annotation tool usually requires trying some of them to know their features and usability. However, the researcher may fail to obtain source or executable code for the many published tools or failing to install the tools due to different technical issues. We discuss some of the available software tools created for image annotation tasks and compare their feature scope in Table 3. Comparison between these tools has been conducted according to seven criteria which are web-based or desktop platform, provide zoom in/out feature, free or need a fee, input types, output types, graphic annotation types and annotation method. For web-based systems, users do not need to download the software on their computers; they just need to open the tools link in a browser when they are ready to annotate the data. In contraposition, if the tool is installed, the user will not need the internet during the annotation process. Manual annotation of images by defining regions of interest by the user is a time-consuming and costly process. Furthermore, it is user dependent. On the other hand, automatic annotation uses ML algorithms trained with pre-annotated images to annotated new sets of images, which is cheaper and faster.

By personally testing all 13 tools above, we found that all of them provide the ability to zoom in and out of specific parts of images. Most of the tools have some limitations not supporting all medical files formats except 3D-Slicer, ITK-SNAP and Supervisely, which accept more types of medical image file extensions to be annotated than the other tools. Another limitation is that some of them are not supported to work on all operating systems for example Reactlabel works only on the MacOS. Also Reactlabel is not free, so its use might be limited. Based on an experiment by Joel et al. [75] on 30 people, the process of importing a set of images into Supervisely was very direct and simple. With Labelbox it was also easy to import images, but it required extra steps such as defining the objects, colour for labels, and the type of tools before getting to the annotating screen. Furthermore, Labelbox does not allow the download of single or specific images but instead requires download of the entire dataset, and the same for Supervisely. In Labelbox, the generated output contains a JSON file, while in Supervisely it may also contain a mask of the annotation in an image format. Joel et al. [75] also found that Supervisely and LableBox are faster than other tools in opening a previously created mask, since both are implemented as cloud based.

LabelIMG tool lacks management properties, such the editing of images [32], which is supported by Supervisely to manage the project in several layers: teams, datasets, and

Table 3 Comparative summary of image annotation tools

Image annotation tool	Platform	Zoom	Free	Input	Output	Graphic annotation	Annotation method
VIA	Web-based	Yes	Yes	Image, video, audio	JSON, CSV, COCO format, png	Polygon, fixed shape	Manual
LabelIMG	Desktop	Yes	Yes	Image	XML files in PASCAL VOC, used format by ImageNet, YOLO	Polygon	Manual
Ratsnake	Desktop	Yes	Yes	Image	Custom text, owl, XML	Polygon, grid, freehand	Manual, automatic
VOTT	Web-based	Yes	Yes	Image, video	Azure Custom Vision Service, CNTK, Pascal VOC, TFRecords, JSON, CSV	Polygon	Manual
Mask Editor	Desktop	Yes	Yes	Image	Binary image files, mat	Polygon, grid, freehand	Manual
Supervisely	Web-based	Yes	Yes	Image, video	JSON, png, TFRecords and XML	Boxes, lines, dots, polygons, bitmap brush	Manual, automatic
RectLabel	Desktop	Yes	No	Image	YOLO, Create ML, COCO, JSON, and CSV formats	Bounding box, polygon, cubic bezier, line and brush	Manual, automatic
LabelMe	Web-based	Yes	Yes	Image	XML files	Polygon, bounding box	Manual
Labelme	Desktop	Yes	Yes	Image, video	JSON file, COCO, VOC format	Polygon, rectangle, circle, line and point	Manual
CVAT	Web-based	Yes	Yes	Image, video	CVAT for video, CVAT for images, PASCAL VOC, (VOC) Segmentation mask, YOLO, COCO, TFRecord, MOT, LabelMe 3.0, Datamaro	Polygon, bounding box, poly-line, points	Manual, automatic
ITK-Snap	Desktop	Yes	Yes	3D Image	DICOM, NIFTI, MetalImage and Analyze	Polygon, brush	Manual, automatic
LabelBox	Web-based	Yes	Yes	Image, video	CSV, JSON	Polygon, bounding box, poly-line, point	Manual
3D-Slicer	Desktop	Yes	Yes	2D and 3D Image	NRRD or NIFTI file	Paint, draw	Manual, automatic

Table 4 The annotation time and cost for each reviewed tool. The first column contains the time in seconds for segmenting the breast arterial calcification in a single mammogram, while the second column show the tool's size in MegaBytes

Tool	Time in second	Tool's size in MegaBytes (MB)
VIA	46.372	N/A
LabelIMG	19.589	8.6 MB
Ratsnake	41.254	57.8 MB
VOTT	44.260	467.8 MB
Mask Editor	48.390	23.8 MB
Supervisely	43.030	N/A
RectLabel	52.634	47.8 MB
LabelMe	55.786	N/A
Labelme	54.238	32.3 MB
CVAT	43.222	N/A
ITK-Snap	54.690	281.7 MB
LabelBox	44.891	N/A
3D-Slicer	33.950	1250.0 MB

workspaces, which is an advantage over LabelIMG. According to Dondi et al. [39] VIA does not require previous technical knowledge. It supports manual annotation requirements and allows users to produce metadata. Most non-technical users prefer to use VIA as it does not require a setup procedure or additional software installation. Moreover, the VIA web-based tool runs easily on all platforms such as Windows, MacOS and Linux. Mallisery et al. [97] used VIA and saved annotations as JSON file. Bernal et al. [16] utilized VIA to enable the user to get the CSV files with the entire annotation results *i.e.*, annotation mask and text metadata values. Ratsnake relies on the use of a snake and directly provides the binary GT masks as a result. It has the appropriate pixel-wise level, is easy to use, and is an open-source annotation tool [52]. Both VOTT and CVAT, focus on single-annotator use instead of allowing many annotators to work simultaneously on the same file [95].

We also evaluated the annotation time and cost for all reviewed tools, which include the segmentation time per image on a 2 GHz Quad-Core Intel Core i5 device. We defined a test case to segment the breast arterial calcification (BAC) from a single mammogram image with size of 2457×1996 pixels. Note that, using the same segmentation method for all the tools was not feasible since they were developed for different applications. Thus, a polygon segmentation method was used in all tools except in LabelIMG, where we used a detection box, and in 3D-Slicer where we used a paint effect. Each tool's time complexity and cost-effectiveness are summarized in Table 4 where N/A refers to that tool web-based that does not require installation. As we can see from the table, between the tools that offer the polygon method, Ratsnake was the fastest one and LabelMe was the slowest one. However, considering the tools' size, the web-based tools (including VIA, Supervisely, CVAT and LabelBox) are more preferable for medical images since they do not need memory space.

Most of the annotation tools rely on semantic web technologies and use metadata attached to the annotated objects. Annotating a particular object in the image using different shapes is performed according to the model's needs to carry out specific operations in the region. Several existing manual annotation tools demand installation and setup procedures. This condition often leads to barriers for non-technical users who cannot deal with setup and installation procedures on the different platforms [100]. User experience design (UXD) is an open research area for annotation tools to enhance usability and accessibility by adding features that increase user satisfaction.

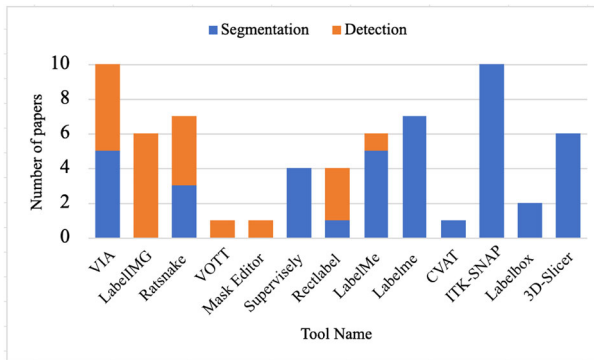


Fig. 22 The papers that used the reviewed tools in term of detection and segmentation task

As observed in Fig. 22 the most used tool in medical task is ITK-SNAP and VIA. However, the VIA tool is used for both segmentation and detection tasks, while the ITK-SNAP tool is employed only for medical image segmentation purposes.

6 Conclusions

Due to the large number of medical images generated, annotation tools software became necessary to tag on these images to improve the medical applications. In this survey, we reviewed various tools that had been used for medical image annotation under several constraints like ease to install and whether they support manual annotation. We presented a brief description and a snapshot for each tool and evaluated them according to different criteria. Also, we mentioned papers that used these tools in different tasks for medical application. Choosing the suitable software tool for annotating the images is an extremely important step that will significantly reduce the amount of work and time needed for annotation. Familiarity with all common tools and decide the suitable one among them is intricate. The aim of this survey is to help the researchers to select the appropriate annotation tool for their task and take a glance at the tool's GUI to simplify dealing with it.

Our study shows that manual image annotation tools are more preferable tools in the medical domain. Manually tagging a large dataset of medical images is considered to be a time-consuming and computationally expensive process, since it needs multiple experts' opinions to avoid human errors. On the other hand, semi-automatic and automatic annotation can achieve better results and save time and effort. Besides, we found that most users prefer web-based tools, which are easy to access and use, and do not require previous technical knowledge. Overall, we hope that this survey will increase awareness of what already exists and reduce resources for expensive functionality developments that are already available. A greater focus on the benefit of deep learning-based models to build automatic medical image annotation tools software will achieve better results. A further study focusing on user requirements is suggested to help the developers identify weaknesses in their tools or for those planning to develop new annotation tools.

References

1. Abtahi S, Omidyeganeh M, Shirmohammadi S, Hariri B (2014) YawDD: Yawning detection dataset. In: Proceedings of the 5th ACM Multimedia Systems Conference, pp 24–28
2. Afouras T, Owens A, Chung JS, Zisserman A (2020) Self-supervised learning of audio-visual objects from video. arXiv:2008.04237
3. Ahmad HA, Yu HJ, Miller CG (2014) Medical imaging modalities. In: Medical Imaging in Clinical Trials. Springer, pp 3–26
4. Al-sudani AR (2020) Yawn based driver fatigue level prediction. Proceedings of 35th International Confer 69:372–382
5. AlGhamdi M, Abdel-Mottaleb M, Collado-Mesa F (2020) DU-Net: Convolutional network for the detection of arterial calcifications in mammograms. IEEE Trans Med Imaging 39(10):3240–3249. <https://doi.org/10.1109/TMI.2020.2989737>
6. Ali S, Zhou F, Daul C, Braden B, Bailey A, Realdon S, East J, Wagnieres G, Loschenov V, Grisan E et al (2019) Endoscopy artifact detection (EAD 2019) challenge dataset. arXiv:1905.03209
7. Ali S, Zhou F, Daul C, Braden B, Bailey A, Realdon S, East J, Wagnieres G, Loschenov V, Grisan E et al (2019) Endoscopy artifact detection (EAD 2019) challenge dataset. arXiv:1905.03209
8. Amini Z, Rabbani H (2016) Classification of medical image modeling methods: A review. Current Medical Imaging 12(2):130–148
9. Aote SS, Potnurwar A (2019) An automatic video annotation framework based on two level keyframe extraction mechanism. Multimedia Tools and Applications 78(11):14465–14484
10. Araújo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, Polónia A, Campilho A (2017) Classification of breast cancer histology images using convolutional neural networks. PLoS one 12(6):e0177544
11. Aumann S, Donner S, Fischer J, Müller F (2019) Optical coherence tomography (OCT): Principle and technical realization. In: High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics. Springer International Publishing, Cham, pp 59–85
12. Bain M, Nagrani A, Schofield D, Zisserman A (2019) Count, crop and recognise: Fine-grained recognition in the wild. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp 0–0
13. Bansal M, Kumar M, Kumar M, Kumar K (2021) An efficient technique for object recognition using shi-tomasi corner detection algorithm. Soft Comput 25(6):4423–4432
14. Barrile V, Candela G, Fotia A (2019) Point cloud segmentation using image processing techniques for structural analysis. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences
15. Belko A, Dobratulin K, Kuznetsov A (2020) Feathers dataset for fine-grained visual categorization. arXiv:2004.08606
16. Bernal J, Histace A, Masana M, Angermann Q, Sánchez-Montes C, de Miguel CR, Hammami M, García-Rodríguez A, Córdova H, Romain O et al (2019) GTCreator: a flexible annotation tool for image-based datasets. Int J CARS 14(2):191–201
17. Besson FL, Henry T, Meyer C, Chevance V, Roblot V, Blanchet E, Arnould V, Grimon G, Chekroun M, Mabile L et al (2018) Rapid contour-based segmentation for 18F-FDG PET imaging of lung tumors by using ITK-SNAP: comparison to expert-based segmentation. Radiology 288(1):277–284
18. Betti A, Michelozzi B, Bracci A, Masini A (2020) Real-time target detection in maritime scenarios based on YOLOv3 model. arXiv:2003.00800
19. Bianco S, Ciocca G, Napoletano P, Schettini R (2015) An interactive tool for manual, semi-automatic and automatic video annotation. Comput Vis Image Underst 131:88–99
20. Biresaw TA, Nawaz T, Ferryman J, Dell AI (2016) Vitbat: Video tracking and behavior annotation tool. In: 2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), IEEE, pp 295–301
21. Brehar R, Mitrea D-A, Vancea F, Marita T, Nedeveschi S, Lupsor-Platon M, Rotaru M, Badea RI (2020) Comparison of deep-learning and conventional machine-learning methods for the automatic recognition of the hepatocellular carcinoma areas from ultrasound images. Sensors 20(11):3085
22. Bromiley PA, Schunke AC, Ragheb H, Thacker NA, Tautz D (2014) Semi-automatic landmark point annotation for geometric morphometrics. Frontiers in Zoology 11(1):61
23. Candemir S, Jaeger S, Palaniappan K, Musco JP, Singh RK, Xue Z, Karargyris A, Antani S, Thoma G, McDonald CJ (2013) Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE transactions on medical imaging 33(2):577–590
24. Cassidy B, Reeves ND, Joseph P, Gillespie D, O’Shea C, Rajbhandari S, Maiya AG, Frank E, Boulton A, Armstrong D et al (2020) DFUC2020: Analysis towards diabetic foot ulcer detection. arXiv:2004.11853

25. Chaichulee S, Villarroel M, Jorge J, Arteta C, Green G, McCormick K, Zisserman A, Tarassenko L (2018) localised photoplethysmography imaging for heart rate estimation of pre-term infants in the clinic. In: *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*, vol 10501, International Society for Optics and Photonics, p 105010R
26. Chen J, Chen L, Wang S, Chen P (2020) A novel multi-scale adversarial networks for precise segmentation of X-ray breast mass. *IEEE Access* 8:103772–103781
27. Chen Y, Wang Y, Hu F, Wang D (2020) A lung dense deep convolutional neural network for robust lung parenchyma segmentation. *IEEE Access* 8:93527–93547
28. Chen Y (2019) Estimating plant phenotypic traits from RGB imagery. Ph.D. Thesis, Purdue University Graduate School
29. Choi B-K, Madusanka N, Choi H-K, So J-H, Kim C-H, Park H-G, Bhattacharjee S, Prakash D (2020) Convolutional neural network-based mr image analysis for alzheimer's disease classification. *Current Medical Imaging* 16(1):27–35
30. Christensen JH, Mogensen LV, Ravn O (2020) Deep learning based segmentation of fish in noisy forward looking MBES images. [arXiv:2006.09034](https://arxiv.org/abs/2006.09034)
31. Ciaparrone G, Bardozzo F, Priscoli MD, Londoño Kallewaard J, Zuluaga MR, Tagliaferri R (2020) a comparative analysis of multi-backbone mask R-CNN for surgical tools detection. In: *2020 International Joint Conference on Neural Networks (IJCNN)*, IEEE, pp 1–8
32. da Silva JL, Tabata AN, Broto LC, Cocron MP, Zimmer A, Brandmeier T (2020) Open source multi-purpose multimedia annotation tool. In: *International Conference on Image Analysis and Recognition*, Springer, pp 356–367
33. Dargan S, Kumar M (2020) A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Syst Appl* 143:113114
34. Dargan S, Kumar M, Ayyagari MR, Kumar G (2019) A survey of deep learning and its applications: a new paradigm to machine learning. *Archives of Computational Methods in Engineering*, pp 1–22
35. Dasiopoulou S, Giannakidou E, Litos G, Malasioti P, Kompatsiaris Y (2011) A survey of semantic image and video annotation tools. In: *Knowledge-driven multimedia information extraction and ontology evolution*. Springer, pp 196–239
36. Deeba F, Mohammed SK, Bui FM, Wahid KA (2017) Efficacy evaluation of save for the diagnosis of superficial neoplastic lesion. *IEEE journal of translational engineering in health and medicine* 5:1–12
37. Dhieb N, Ghazzai H, Besbes H, Massoud Y (2019) An automated blood cells counting and classification framework using mask R-CNN deep learning model. In: *2019 31st International Conference on Microelectronics (ICM)*, IEEE, pp 300–303
38. Dias PA, Shen Z, Tabb A, Medeiros H (2019) FreeLabel: A publicly available annotation tool based on freehand traces. In: *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, pp 21–30
39. Dondi C, Dutta A, Malaspina M, Zisserman A (2020) The use and reuse of printed illustrations in 15th-century venetian editions. *Printing R-Evolution and Society*
40. Dong X, Yan Y, Ouyang W, Yang Y (2018) Style aggregated network for facial landmark detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 379–388
41. Dou Q, Chen H, Jin Y, Yu L, Qin J, Heng P-A (2016) 3D deeply supervised network for automatic liver segmentation from CT volumes. In: *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Springer, pp 149–157
42. Dutta A, Zisserman A (2019) The via annotation software for images, audio and video. In: *proceedings of the 27th acm international conference on multimedia*, pp 2276–2279
43. Ezhilarasi R, Varalakshmi P (2018) Tumor detection in the brain using faster R-CNN. In: *2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, 2018 2nd International Conference on, IEEE, pp 388–392
44. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M et al (2012) 3D slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging* 30(9):1323–1341
45. Ferlaino M, Glastonbury CA, Motta-Mejia C, Vatish M, Granne I, Kennedy S, Lindgren CM, Nellåker C (2018) Towards deep cellular phenotyping in placental histology. [arXiv:1804.03270](https://arxiv.org/abs/1804.03270)
46. François T, Calvet L, Madad S, Saboul D, Gasparini S, Samarakoon P, Bourdel N, Bartoli A (2020) Detecting the occluding contours of the uterus to automatise augmented laparoscopy: score, loss, dataset, evaluation and user study. *International journal of computer assisted radiology and surgery*
47. Fukuda M, Okuno T, Yuki S (2020) Central object segmentation by deep learning for fruits and other roundish objects. [arXiv:2008.01251](https://arxiv.org/abs/2008.01251)

48. Gambella C, Ghaddar B, Naoum-Sawaya J (2020) Optimization problems for machine learning: a survey. *Eur J Oper Res*
49. Gaonkar B, Edwards M, Bui A, Brown M, Macyszyn L (2018) Extreme augmentation: Can deep learning based medical image segmentation be trained using a single manually delineated scan?. arXiv:1810.01621
50. Gaur E, Saxena V, Singh SK (2018) Video annotation tools: A review. In: 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCCN), pp 911–914
51. Gentil M, Sameki M, Gurari D, Saraee E, Hasenberg E, Wong JY, Betke M (2016) Interactive tracking of cells in microscopy image sequences. In: Proceedings of the Third Interactive Medical Image Computation Workshop (IMIC) at The Medical Image Computing and Computer Assisted Intervention Society (MICCAI)
52. Ghanem S, Imran A, Athitsos V (2019) Analysis of hand segmentation on challenging hand over face scenario. In: Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments, pp 236–242
53. Ghosh KK, Begum S, Sardar A, Adhikary S, Ghosh M, Kumar M, Sarkar R (2021) Theoretical and empirical analysis of filter ranking methods: Experimental study on benchmark dna microarray data. *Expert Syst Appl* 169:114485
54. Gillespie D, Yap MH, Hewitt BM, Driscoll H, Simanaviciute U, Hodson-Tole EF, Grant RA (2019) Description and validation of the LocoWhisk system: Quantifying rodent exploratory, sensory and motor behaviours. *J Neurosci Methods* 328:108440
55. Gou M, Rao Y, Zhang M, Sun J, Cheng K (2019) Automatic image annotation and deep learning for tooth CT image segmentation. In: International Conference on Image and Graphics, Springer, pp 519–528
56. Gupta S, Thakur K, Kumar M (2020) 2d-human face recognition using sift and surf descriptors of face?s feature regions. *Vis Comput*, pp 1–10
57. Gupte T, Niljkar M, Gawali M, Kulkarni V, Kharat A, Pant A (2021) Deep learning models for calculation of cardiothoracic ratio from chest radiographs for assisted diagnosis of cardiomegaly. arXiv:2101.07606
58. Gurari D, Theriault D, Sameki M, Isenberg B, Pham TA, Purwada A, Solski P, Walker M, Zhang C, Wong JY et al (2015) How to collect segmentations for biomedical images? a benchmark evaluating the performance of experts, crowdsourced non-experts, and algorithms. In: 2015 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, pp 1169–1176
59. Hadush S, Girmay Y, Sinamo A, Hagos G (2020) Breast cancer detection using convolutional neural networks. arXiv:2003.07911
60. Hahn S, Morris CS, Bertges DJ, Wshah S (2019) Deep learning for recognition of endoleak after endovascular abdominal aortic aneurysm repair. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, pp 759–763
61. Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P (2000) The digital database for screening mammography. Proceedings of the Fourth International Workshop on Digital Mammography. https://doi.org/10.1007/978-94-011-5318-8_75
62. Hidayatullah P, Mengko TER, Munir R, Barlian A (2019) A semiautomatic sperm cell data annotator for convolutional neural network. In: 2019 5th International Conference on Science in Information Technology (ICSITech), IEEE, pp 211–216
63. Hong J, Fulton M, Sattar J (2020) TrashCan: A semantically-segmented dataset towards visual detection of marine debris. arXiv:2007.08097
64. Hosseini SMHadi, Chen H, Jablonski MM (2020) Automatic detection and counting of retina cell nuclei using deep learning. In: Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol 11317, International Society for Optics and Photonics, p 113172I
65. Hu K, Chen K, He X, Zhang Y, Chen Z, Li X, Gao X (2020) Automatic segmentation of intracerebral hemorrhage in CT images using encoder–decoder convolutional neural network. *Information Processing & Management* 57(6):102352
66. Huang J, Shen H, Wu J, Hu X, Zhu Z, Lv X, Liu Y, Wang Y (2020) Spine Explorer: a deep learning based fully automated program for efficient and reliable quantifications of the vertebrae and discs on sagittal lumbar spine mr images. *The Spine Journal* 20(4):590–599
67. Iakovidis D, Goudas T, Smailis C, Maglogiannis I (2014) Ratsnake: a versatile image annotation tool with application to computer-aided diagnosis. *The Scientific World Journal* 2014:286856. <https://doi.org/10.1155/2014/286856>
68. Iakovidis DK, Chatzis D, Chrysanthopoulos P, Koulaouzidis A (2015) Blood detection in wireless capsule endoscope images based on salient superpixels. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp 731–734

69. Iakovidis DK, Koulaouzidis A (2014) Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software. *Gastrointest Endosc* 80(5):877–883. <https://doi.org/10.1016/j.gie.2014.06.026>, <http://www.sciencedirect.com/science/article/pii/S0016510714018616>
70. Iakovidis DK, Koulaouzidis A (2014) Automatic lesion detection in wireless capsule endoscopy? a simple solution for a complex problem. In: 2014 IEEE International Conference on Image Processing (ICIP), IEEE, pp 2236–2240
71. Igloukov VI, Rakhlin A, Kalinin AA, Shvets AA (2018) Paediatric bone age assessment using deep convolutional neural networks. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, pp 300–308
72. Intel (2018) Computer Vision Annotation Tool (CVAT). <https://github.com/openvinotoolkit/cvat>
73. Jantsho Y, Riyamongkol P, Waranusast R (2020) Real-time bhutanese license plate localization using YOLO. *ICT Express* 6(2):121–124
74. Jha D, Smedsrud PH, Riegler MA, Halvorsen P, de Lange T, Johansen D, Johansen HD (2020) Kvasir-SEG: A segmented polyp dataset. In: *International Conference on Multimedia Modeling*, Springer, pp 451–462
75. Joel B-G, Hellen R-M, Adrián G-A, Saúl C-R, Fabian P-J, Carlos C-AL, Ricardo B-C (2019) Insight GT: A public, fast, web image ground truth authoring tool. In: *Latin American High Performance Computing Conference*, Springer, pp 398–405
76. Kasban H, El-Bendary MAM, Salama DH (2015) A comparative study of medical imaging techniques. *International Journal of Information Science and Intelligent System* 4(2):37–58
77. Kaur P, Kumar R, Kumar M (2019) A healthcare monitoring system using random forest and internet of things (iot). *Multimedia Tools and Applications* 78(14):19905–19916
78. Kawamura R (2017) Rectlabel application for annotation. <https://rectlabel.com/>
79. Kawazoe Y, Shimamoto K, Yamaguchi R, Shintani-Domoto Y, Uozaki H, Fukayama M, Ohe K (2018) Faster R-CNN-based glomerular detection in multistained human whole slide images. *Journal of Imaging* 4(7):91
80. Khaki S, Pham H, Han Y, Kuhl A, Kent W, Wang L (2020) Convolutional neural networks for image-based corn kernel detection and counting. *Sensors* 20(9):2721
81. Kim JY, Ro K, You S, Nam BR, Yook S, Park HS, Yoo JC, Park E, Cho K, Cho BH et al (2019) Development of an automatic muscle atrophy measuring algorithm to calculate the ratio of supraspinatus in supraspinous fossa using deep learning. *Comput Methods Prog Biomed* 182:105063
82. Kondal S, Kulkarni V, Gaikwad A, Kharat A, Pant A (2020) Automatic grading of knee osteoarthritis on the Kellgren-Lawrence scale from radiographs using convolutional neural networks. arXiv:2004.08572
83. Kordon F, Maier A, Swartman B, Privalov M, El Barbari JS, Kunze H (2020) Contour-based bone axis detection for X-ray guided surgery on the knee. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp 671–680
84. Koulaouzidis A, Iakovidis DK, Yung DE, Rondonotti E, Kopylov U, Plevris JN, Toth E, Eliakim A, Johansson GW, Marlicz W et al (2017) KID project: an internet-based digital video atlas of capsule endoscopy for research purposes. *Endoscopy international open* 5(6):E477
85. Kumar A, Kumar M, Kaur A (2021) Face detection in still images under occlusion and non-uniform illumination. *Multimedia Tools and Applications* 80(10):14565–14590
86. Kumar M, Gupta S, Kumar K, Sachdeva M (2020) Spreading of covid-19 in india, italy, japan, spain, uk, us: a prediction using arima and lstm model. *Digital Government: Research and Practice* 1(4):1–9
87. Kummerfeld JK (2019) SLATE: a super-lightweight annotation tool for experts. arXiv:1907.08236
88. Larobina M, Murino L (2014) Medical image file formats. *J Digit Imaging* 27(2):200–206
89. Lee E-J, Plishker W, Liu X, Bhattacharyya SS, Shekhar R (2019) Weakly supervised segmentation for real-time surgical tool tracking. *Healthcare Technology Letters* 6(6):231–236
90. Lee SK (2020) Pig pose estimation based on extracted data of mask R-CNN with VGG neural network for classifications. Master's Thesis, South Dakota State University
91. Li C, Zhang D, Chen S (2020) Research about tongue image of traditional chinese medicine (TCM) based on artificial intelligence technology. In: *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, IEEE, pp 633–636
92. Li Z, Wang C, Han M, Xue Y, Wei W, Li L-J, Fei-Fei L (2018) Thoracic disease identification and localization with limited supervision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 8290–8299
93. Liu F, Liu D, Tian J, Xie X, Yang X, Wang K (2020) Cascaded one-shot deformable convolutional neural networks: Developing a deep learning model for respiratory motion estimation in ultrasound sequences. *Med Image Anal* 65:101793

94. Liu Y, Wang J, Zhong S (2020) Correlation between clinical risk factors and tracheal intubation difficulty in infants with Pierre-Robin syndrome: a retrospective study. *BMC anesthesiology* 20(1):1–6
95. Lynnette NHX, Hock HNS, Yen NY (2020) Cross-model image annotation platform with active learning. arXiv:2008.02421
96. Macarini LAB, von Wangenheim A, Daltoé FP, Onofre ASC, Onofre FBM, Stemmer MR (2020) Towards a complete pipeline for segmenting nuclei in feulgen-stained images. arXiv:2002.08331
97. Malliserry S, Wu M-C, Bau C-A, Huang G-Z, Yang C-Y, Lin W-C, Wu Y-S (2020) POSTER: Data leakage detection for health information system based on memory introspection. In: Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, pp 898–900
98. Mansoor A, Cerrolaza JJ, Idrees R, Biggs E, Alsharid MA, Avery RA, Linguraru MG (2016) Deep learning guided partitioned shape model for anterior visual pathway segmentation. *IEEE transactions on medical imaging* 35(8):1856–1865
99. Marzahl C, Aubreville M, Bertram CA, Gerlach S, Maier J, Voigt J, Hill J, Klopfeisch R, Maier A (2019) Fooling the crowd with deep learning-based methods. arXiv:1912.00142
100. Microsoft (2019) VoTT (Visual Object Tagging Tool). <https://github.com/microsoft/VoTT>
101. Miok K, Pirs G, Robnik-Sikonja M (2020) Bayesian methods for semi-supervised text annotation. arXiv:2010.14872
102. Moehrmann J, Heidemann G (2012) Efficient annotation of image data sets for computer vision applications. In: Proceedings of the 1st International Workshop on Visual Interfaces for Ground Truth Collection in Computer Vision Applications, pp 1–6
103. Müller M, Dohrn MF, Romanzetti S, Gadermayr M, Reetz K, Krämer NA, Kuhl C, Schulz JB, Gess B (2020) Semi-automated volumetry of MRI serves as a biomarker in neuromuscular patients. *Muscle & nerve* 61(5):600–607
104. Nakasi R, Mwebaze E, Zawedde A, Tusubira J, Akera B, Maiga G (2020) A new approach for microscopic diagnosis of malaria parasites in thick blood smears using pre-trained deep learning models. *SN Applied Sciences* 2(7):1–7
105. Neves M, Leser U (2014) A survey on annotation tools for the biomedical literature. *Briefings in bioinformatics* 15(2):327–340
106. Nobis F, Geisslinger M, Weber M, Betz J, Lienkamp M (2019) A deep learning-based radar and camera sensor fusion architecture for object detection. In: 2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE, pp 1–7
107. Nowak S, Rüger S (2010) How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation. In: Proceedings of the international conference on Multimedia information retrieval, pp 557–566
108. Ohee MNS, Asif M (2020) real-time tiger detection using YOLOv3. *International Journal of Computer Applications* 975:8887
109. Park A, Chute C, Rajpurkar P, Lou J, Ball RL, Shpanskaya K, Jabarkheel R, Kim LH, McKenna E, Tseng J et al (2019) Deep learning–assisted diagnosis of cerebral aneurysms using the HeadXNet model. *JAMA network open* 2(6):e195600–e195600
110. Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans Med Imaging* 35(5):1240–1251
111. Prado E, Rodríguez-Basalo A, Cobo A, Ríos P, Sánchez F (2020) 3D fine-scale terrain variables from underwater photogrammetry: A new approach to benthic microhabitat modeling in a circalittoral rocky shelf. *Remote Sens* 12(15):2466
112. Rahim T, Hassan SA, Shin SY (2020) A deep convolutional neural network for the detection of polyps in colonoscopy images. arXiv:2008.06721
113. Rajaraman S, Sornapudi S, Alderson PO, Folio LR, Antani SK (2020) Analyzing inter-reader variability affecting deep ensemble learning for COVID-19 detection in chest radiographs. *PloS one* 15(11):e0242301
114. Rasoulia A, Rohling RN, Abolmaesumi P (2013) A statistical multi-vertebrae shape+ pose model for segmentation of CT images. In: Medical Imaging 2013: Image-Guided Procedures, Robotic Interventions, and Modeling, vol 8671, International Society for Optics and Photonics, p 86710P
115. Ray S (2019) A quick review of machine learning algorithms. In: 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon), IEEE, pp 35–39
116. Rebinth A, Kumar S M (2019) Importance of manual image annotation tools and free datasets for medical research. *Journal of Advanced Research in Dynamical and Control Systems* 10:1880–1885
117. Rella S (2020) Distributed collaborative framework for deep learning in object detection. Master's Thesis
118. Roihan A, Hasanudin M, Sunandar E (2020) Evaluation methods of bird repellent devices in optimizing crop production in agriculture. In: *Journal of Physics: Conference Series*, vol 1477, p 032012

119. Roth HR, Lu L, Farag A, Shin H-C, Liu J, Turkbey EB, Summers RM (2015) Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Springer, pp 556–564
120. Roth HR, Yang D, Xu Z, Wang X, Xu D (2020) Going to extremes: Weakly supervised medical image segmentation. arXiv:2009.11988
121. Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, Huijben I, Chennakeshava N, Mento F, Sentelli A et al (2020) Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. *IEEE Trans Med Imaging*
122. Rubin DL, Akdogan MU, Altindag C, Alkim E (2019) ePAD: An image annotation and analysis platform for quantitative imaging. *Tomography* 5(1):170
123. Russell BC, Torralba A, Murphy KP, Freeman WT (2008) LabelMe: a database and web-based tool for image annotation. *International journal of computer vision* 77(1-3):157–173
124. Rusu M, Rajiah P, Gilkeson R, Yang M, Donatelli C, Thawani R, Jacono FJ, Linden P, Madabhushi A (2017) Co-registration of pre-operative CT with ex vivo surgically excised ground glass nodules to define spatial extent of invasive adenocarcinoma on in vivo imaging: a proof-of-concept study. *European radiology* 27(10):4209–4217
125. Sánchez JCG, Magnusson M, Sandborg M, Tedgren A, AsaCarlsson, Malusek A (2020) Segmentation of bones in medical dual-energy computed tomography volumes using the 3D U-Net. *Physica Medica* 69:241–247
126. Sha G, Wu J, Yu B (2020) Detection of spinal fracture lesions based on improved YOLOv2. In: 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), IEEE, pp 235–238
127. Sharma M, Rasmuson D, Rieger B, Kjelkerud D et al (2019) Labelbox: The best way to create and manage training data. software, labelbox. Inc. <https://www.labelbox.com>
128. Shen J, Baum T, Cordes C, Ott B, Skurk T, Kooijman H, Rummeny EJ, Hauner H, Menze BH, Karapinos DC (2016) Automatic segmentation of abdominal organs and adipose tissue compartments in water-fat MRI: application to weight-loss in obesity. *European journal of radiology* 85(9):1613–1621
129. Siam M, Jiang C, Lu S, Petrich L, Gamal M, Elhoseiny M, Jagersand M (2019) Video object segmentation using teacher-student adaptation in a human robot interaction (HRI) setting. In: 2019 International Conference on Robotics and Automation (ICRA), IEEE, pp 50–56
130. Singh PP, Prasad S, Chaudhary AK, Patel CK, Debnath M (2019) Classification of effusion and cartilage erosion affects in osteoarthritis knee MRI images using deep learning model. In: International Conference on Computer Vision and Image Processing, Springer, pp 373–383
131. Singh S, Ahuja U, Kumar M, Kumar K, Sachdeva M (2021) Face mask detection using yolov3 and faster r-cnn models: Covid-19 environment. *Multimedia Tools and Applications* 80(13):19753–19768
132. Sirazitdinov I, Schulz H, Saalbach A, Renisch S, Dylov DV (2020) Tubular shape aware data generation for semantic segmentation in medical imaging. arXiv:2010.00907
133. Song T, Meng F, Rodriguez-Paton A, Li P, Zheng P, Wang X (2019) U-Next: A novel convolution neural network with an aggregation U-Net architecture for gallstone segmentation in CT images. *IEEE Access* 7:166823–166832
134. Staal J, Abràmoff MD, Niemeijer M, Viergever MA, Van Ginneken B (2004) Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging* 23(4):501–509
135. Sun C, Guo S, Zhang H, Li J, Chen M, Ma S, Jin L, Liu X, Li X, Qian X (2017) Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs. *Artificial intelligence in medicine* 83:58–66
136. Systems D (2017) supervisely web platform for computer vision. annotation, training and deploy. <https://supervisely/>
137. Tang H, Sun N, Li Y (2020) Segmentation model of the opacity regions in the chest X-rays of the COVID-19 patients in the us rural areas and the application to the disease severity. medRxiv
138. Tolkachev A, Sirazitdinov I, Kholiavchenko M, Mustafaev T, Ibragimov B (2020) Deep learning for diagnosis and segmentation of pneumothorax: The results on the Kaggle competition and validation against radiologists. *IEEE Journal of Biomedical and Health Informatics*
139. Ullah H, Uzair M, Ullah M, Khan A, Ahmad A, Khan W (2017) Density independent hydrodynamics model for crowd coherency detection. *Neurocomputing* 242:28–39
140. Vats V, Goel P, Agarwal A, Goel N (2020) SURF-SVM based identification and classification of gastrointestinal diseases in wireless capsule endoscopy. arXiv:2009.01179
141. Vickery S, Hopkins WD, Sherwood CC, Schapiro SJ, Litzman RD, Caspers S, Gaser C, Eickhoff SB, Dahnke R, Hoffstaedter F (2020) Chimpanzee brain morphometry utilizing standardized MRI preprocessing and macroanatomical annotations. *Elife* 9:e60136

142. Vlontzos A, Mikolajczyk K (2018) Deep segmentation and registration in X-ray angiography video. arXiv:1805.06406
143. Vostrikov A, Chernyshev S (2019) Training sample generation software. In: Intelligent Decision Technologies 2019. Springer, pp 145–151
144. Wada K (2016) labelme: Image Polygonal Annotation with Python. <https://github.com/wkentaro/labelme>
145. Wang F, Zhou S, Panev S, Han J, Huang D (2019) Person-in-WiFi: Fine-grained person perception using WiFi. In: Proceedings of the IEEE International Conference on Computer Vision, pp 5452–5461
146. Wei JW, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, Tomita N, Abdollahi B, Kim AS, Snover DC et al (2020) Evaluation of a deep neural network for automated classification of colorectal polyps on histopathologic slides. *JAMA Network Open* 3(4):e203398–e203398
147. Weinstein BG, Marconi S, Bohlman S, Zare A, White E (2019) Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. *Remote Sens* 11(11):1309
148. Weinstein BG, Marconi S, Bohlman SA, Zare A, White EP (2020) Cross-site learning in deep learning RGB tree crown detection. *Ecological Informatics* 56:101061
149. Xian Z, Wang X, Yan S, Yang D, Chen J, Peng C (2020) Main coronary vessel segmentation using deep learning in smart medical. *Math Probl Eng*, 2020
150. Xie M, Li Y, Xue Y, Shafritz R, Rahimi SA, Ady JW, Roshan UW (2019) Vessel lumen segmentation in internal carotid artery ultrasounds with deep convolutional neural networks. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, pp 2393–2398
151. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J et al (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering*
152. Yang J, Zhang Y, Li L, Li X (2017) YEDDA: A lightweight collaborative text span annotation tool. arXiv:1711.03759
153. Yi X, Walia E, Babyn P (2018) Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by wasserstein distance for dermoscopy image classification. arXiv:1804.03700
154. Yu C-W, Chen Y-L, Lee K-F, Chen C-H, Hsiao C-Y (2019) efficient intelligent automatic image annotation method based on machine learning techniques. In: 2019 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), IEEE, pp 1–2
155. Yudin DA, Skrynnik A, Krishtopik A, Belkin I, Panov AI (2019) Object detection with deep neural networks for reinforcement learning in the task of autonomous vehicles path planning at the intersection. *Optical Memory and Neural Networks* 28(4):283–295
156. Yushkevich PA, Gao Y, Gerig G (2016) ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp 3342–3345
157. Yushkevich PA, Pashchinskiy A, Oguz I, Mohan S, Schmitt JE, Stein JM, Zukić D, Vicory J, McCormick M, Yushkevich N et al (2019) User-guided segmentation of multi-modality medical imaging datasets with ITK-SNAP. *Neuroinformatics* 17(1):83–102
158. Zadeh SM, Francois T, Calvet L, Chauvet P, Canis M, Bartoli A, Bourdel N (2020) SurgAI: deep learning for computerized laparoscopic image understanding in gynaecology. *Surgical endoscopy* 34(12):5377–5383
159. Zaki G, Gudla PR, Lee K, Kim J, Ozbun L, Shachar S, Gadkari M, Sun J, Fraser IainDC, Franco LM et al (2020) A deep learning pipeline for nucleus segmentation. *Cytometry Part A* 97(12):1248–1264
160. Zhang C, Loken K, Chen Z, Xiao Z, Kunkel G (2018) Mask editor : an image annotation tool for image segmentation tasks. arXiv:1809.06461
161. Zhang F, Wu S, Zhang C, Chen Q, Yang X, Jiang K, Zheng J (2019) Multi-domain features for reducing false positives in automated detection of clustered microcalcifications in digital breast tomosynthesis. *Medical physics* 46(3):1300–1308
162. Zhu G, Piao Z, Kim SC (2020) Tooth detection and segmentation with mask R-CNN. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC), IEEE, pp 070–072