US-Pro

An Application Enabling Efficient, High-Throughput Ultrasound Video Processing

Adrian D. Haimovich, MD, PhD, Zachary Lehmann, BA, R. Andrew Taylor, MD, MHS

Supplemental material online at jultrasoundmed.org

We describe a new graphical user interface–based application, US-Pro, designed to enable customized, high-throughput ultrasound video anonymization and dynamic cropping before output to video or high-efficiency disk storage. This application is distributed in a Docker container environment, which supports facile software installation on the most commonly used operating systems, as well as local processing of data sets, precluding the external transfer of electronic protected health information. The US-Pro application will facilitate the reproducible production of large-scale ultrasound video data sets for varied applications, including machine-learning analysis, educational distribution, and quality assurance review.

Key Words—informatics; machine learning; research; ultrasound video processing

Ultrasound (US) is an imaging modality that enables point-of-care medical evaluations without exposing patients to ionizing radiation.[1] Ultrasound provides diagnostic value in a multitude of clinical scenarios[2–6] and is increasingly considered a core component of future clinical practice, as evidenced by its introduction into residency and medical school curricula.[7,8] Furthermore, recent technical advances are likely to accelerate adoption of individual point-of-care use of US by reducing costs and device sizes.[9] As US demand and the number and variety of users increase, methods and technology to automate processes and enhance quality will become necessary.

Previous work has established methods for batch cropping US videos based on fixed pixels,[10] but this technique requires user conversion from the Digital Imaging and Communications in Medicine (DICOM) format and calibration for different transducer types and machines. In addition, since this method is built on cropping, it does not address the challenge of text and other data embedded in the middle of the clip. One algorithmic approach to US image anonymization had an anonymization success rate of 89.2% but was not available for testing.[11] This technique is limited by the use of US stills and therefore does not leverage the information contained in contiguous frames. Furthermore, none of the described approaches address the challenge of preparing US for subsequent machine-learning applications.[10,11]

Over the past few years, machine-learning techniques termed "deep learning" have come to prominence in image classification problems.[12,13] Prior studies have demonstrated the ability of deep-

Received June 8, 2018, from the Department of Emergency Medicine, Yale University School of Medicine, New Haven, Connecticut, USA. Manuscript accepted for publication December 24, 2018.

Dr Haimovich is supported by National Institutes of Health grants 1F30CA196191 and T32GM007205.

Address correspondence to R. Andrew Taylor, MD, MHS, Department of Emergency Medicine, Yale University School of Medicine, 464 Congress Ave, Suite 260, New Haven, CT 06519 USA.

E-mail: richard.taylor@yale.edu

Abbreviations
DICOM, Digital Imaging and Communications in Medicine; ePHI, electronic protected health information; GUI, graphical user interface; HDF5, hierarchical data format 5; US, ultrasound

doi:10.1002/jum.14951

learning networks to detect diabetic retinopathy in retinal fundal photographs,[14] classify skin lesions,[15] and diagnose pneumonia on chest radiographs at an expert level.[16] However, there have been few reports extending deep-learning techniques to US.[17,18] One key difference between US data and other radiologic modalities is that the clinical interpretation of US videos often requires the comparison of multiple frames. For example, in echocardiography, myocardial mechanics are assessed as a function of time.[19,20] Although deep learning has been applied successfully to videos,[21] extensions to US remain in the development stages.[20,22] Deep learning requires massive numbers of training samples, necessitating the use of automated preprocessing steps and efficient storage.

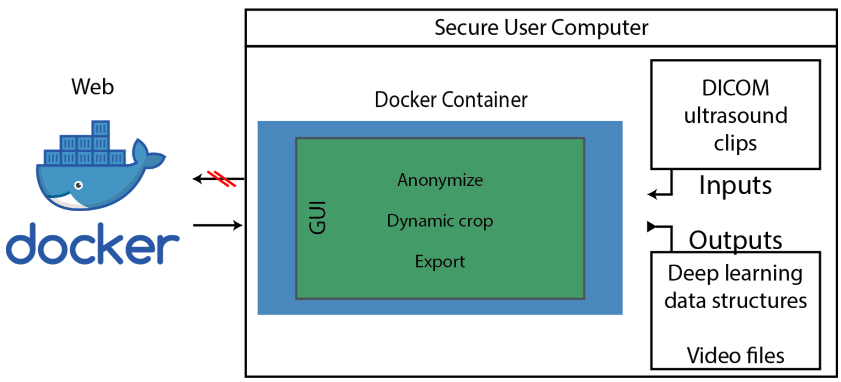
Given the widespread availability of machine-learning algorithms,[23–25] commoditization of computer resources,[26] and success of data crowd-sourcing competitions,[27] there is a strong incentive to mobilize large hospital system-based US video data sets for new clinical and research applications. We present US-Pro, a new graphical user interface (GUI)-based US video-processing application that reads DICOM-standard video files, optionally removes electronic protected health information (ePHI), and exports the data in one of many formats (Figure 1). To facilitate reproducibility and ease of use, we bundle the application into a single Docker container that contains all dependencies and libraries. We test the application on data captured on different US machines with an array of transducers to evaluate the generalizability of our approach. The US-Pro application has been made available to the larger scientific community via GitHub and the Docker image repository (online supplemental "Supporting Information").

Methods

This study was exempted from Institutional Review Board approval by the Yale Human Research Protection Program. Reproducibility and code access are key challenges in the adoption of new computational methods.[28,29] Application installation is managed via Docker; all dependencies and libraries are automatically downloaded (instructions in online supplemental "Supporting Information").[30] Docker containers are conceptually similar to virtual machines but have improved portability, have less computational overhead, and allow dynamic resource allocation.[31] The container includes all packages used, including SimpleITK for video decompression and processing, scikit-learn and numpy for image processing, and pickle, h5py, and skvideo.io for outputs, creating a one-step download for this application that ensures compatibility. All code is available on GitHub (https://github.com/adhaimovich/USPro).

To facilitate pipeline use, we developed a Django-based GUI written in Python that automatically communicates with the processing algorithms (Figure 2).[30,32] This GUI launches automatically when the user runs the container. At run time, the user is required to identify the local system folder, whose subdirectories contain the DICOM video files.

Figure 1. US-Pro deployment pipeline. Users download a publicly available Docker container before launching on host machines. Digital Imaging and Communications in Medicine clips are made available to the Docker container scripts, which process the data locally.



The entirety of the mounted data set is automatically processed as a batch on each execution.

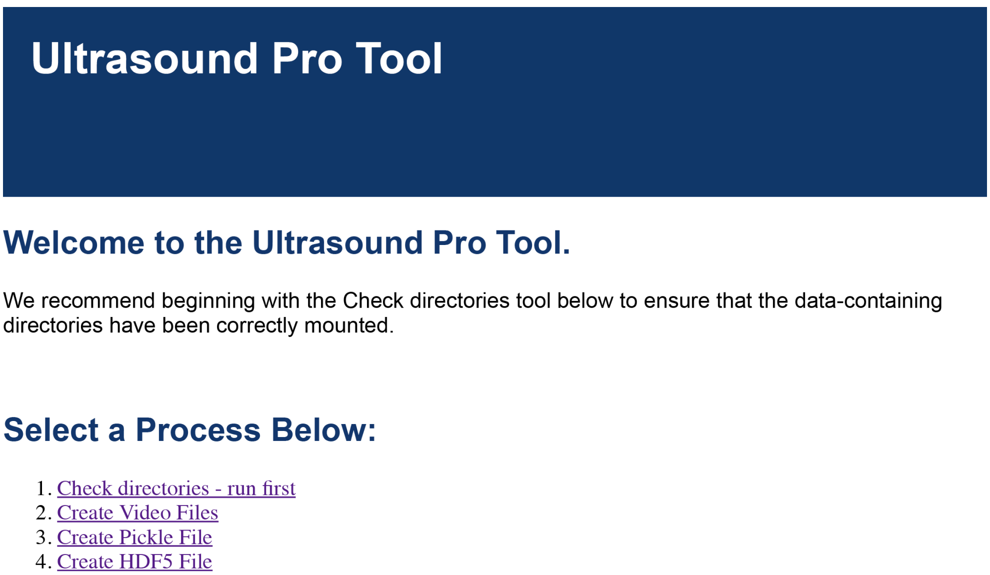
The DICOM clip-processing pipeline begins with decompression using the SimpleITK library, followed by conversion to gray scale. We then use an anonymization step, whose goal is to identify pixels representing common static elements, including image watermarking (eg, instrument model), ePHI (eg, patient identifiers), and scan information (eg, transducer type; Figure 3A). We first calculate the pixel-wise standard deviation over consecutive frames from a single clip. Pixels that are unchanged over the entire clip are reassigned a null value. The resulting cleaned image typically contains the prominent US cone and trace watermarking elements (Figure 3B). The dimensions of the cone are a function of manufacturer design as well as the transducer used in the medical study.[1]

We use image segmentation to dynamically identify the boundaries of the US cone.[33] The image is first binarized by using either a user-defined threshold value or Otsu's method[33] and then segmented by using contiguous pixel labeling (Figure 3C). The largest image segment is determined to be the US cone and used for cropping (Figure 3D). Bilinear interpolation is used to reshape the image to user-defined dimensions. Reducing image dimensions (eg, 600 × 800 to 256 × 256) with down-sampling saves disk space and processes images in preparation for deep-learning frameworks, which typically use smaller images.[34] This step is intended to facilitate transfer learning, ie, the use of pretrained image classification networks, which enables the efficient training of data sets containing fewer samples.[35]

The US-Pro application includes a number of output formats (Figure 1). First, US videos can be exported as video files viewable without a specialized DICOM interpreter to facilitate alternative analysis techniques and anonymized sharing. Alternatively, the US video contents can be exported as a Python-specific pickle file[36] or in hierarchical data format 5 (HDF5),[37] which is a high-performance, flexible data storage format that enables dynamic access while minimizing memory use.[37,38] Of note, both pickle and HDF5 are compatible with downstream deep-learning packages, including TensorFlow[25] and Keras.[23] Furthermore HDF5 data can be stored with lossless compression, including the gzip and Lempel-Ziv-Markov filters.[39] Data exported in this format are automatically reshaped to uniform user-defined dimensions. Descriptions of the output data structures used for these formats are provided in online supplemental "Supporting Information."

An assessment of software performance in anonymization and cropping was performed independently by 2 reviewers. Digital Imaging and Communications in Medicine videos were first converted to audio-video interleave formatting using the US-Pro tool

Figure 2. Software GUI.

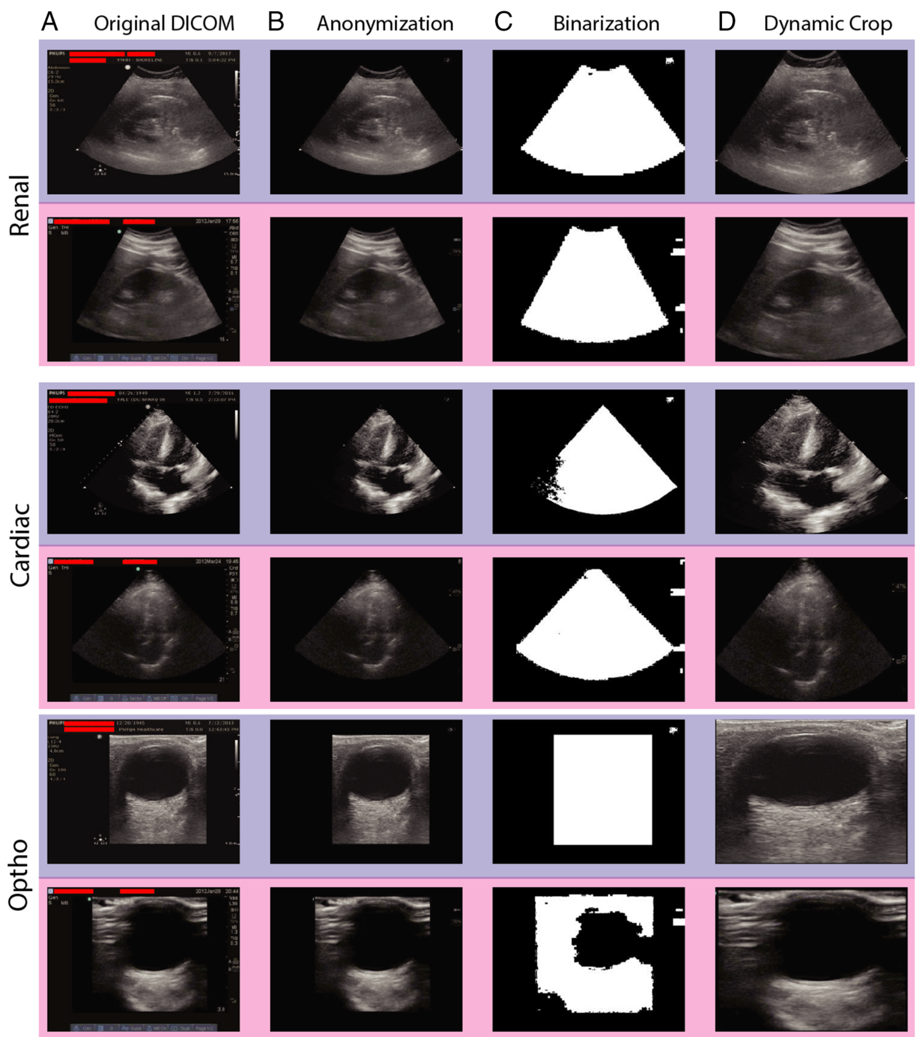


with no processing steps selected. These samples were considered unmodified and used as comparisons for anonymization and cropping. Successful anonymization was defined as complete removal of patient information, including name, date of birth, and medical record number, from all frames. To assess cropping, either a binary threshold value of 1e-07 (estimating 0 but accounting for floating-point calculations) or Otsu's method was used, and the output clips were compared with the negative control. A successful crop was defined as one that included the entirety of the US cone by visual inspection. Inter-reviewer agreement was calculated after reviewing all frames within a video clip using the Cohen κ score in the Python scikit-learn library.[40]

Results

We tested this software on 100 compressed DICOM videos captured during routine clinical care at an academic hospital emergency department and extracted

Figure 3. Ultrasound video-processing pipeline with representative images. Videos from renal, cardiac, and ophthalmologic scans were captured with curvilinear, phased array, and linear transducers, respectively. Ultrasound examinations were performed on instruments from 2 different manufacturers: Philips Healthcare (Bothell, WA; blue boxes) and SonoSite, Inc (Bothell, WA; pink boxes). Images shown were extracted from the first frames of the processed video clips. Red boxes represent redacted patient ePHI.



from a QPath database (Table 1). We observed that 100% of clips were anonymized successfully, deleting ePHI with perfect inter-reviewer agreement ($\kappa = 1$). We also noted that anonymization also removed a number of other image elements, including the manufacturer label, depth markers, and user-entered image labels. For the binary and Otsu thresholding, we observed 96% ($\kappa = 1$) and 14.5% ($\kappa = 0.87$) successful crops, respectively. Given the limited utility of Otsu thresholding, we do not include the technique in the final software package. We then evaluated the 4 clips that failed binary cropping, observing that all were captured by a phased-array transducer on a Philips Sparq machine during cardiac examinations. Two clips were captured in the parasternal long-axis view, and 2 represented the parasternal short-axis view. We subsequently extracted 20 more clips captured during routine clinical care that were evenly divided between these views and analyzed them with this pipeline. Consistent with our earlier results, we observed 100% anonymization ($\kappa = 1$). All 20 clips were correctly cropped by binary thresholding ($\kappa = 1$).

Since our approach leverages an image segmentation step, we then calculated the number of pixels within the calculated US cone that were blanked during the initial anonymization step. For this analysis, we used only videos in the original data set ($n = 96$) that were cropped correctly by the binary threshold method. We found a mean and SD of 0.21% and 0.54% pixels modified, respectively. Only 5 clips had greater than 1% of pixels modified. All 5 clips were lung windows captured with a linear transducer. On manual review, we noted that these images had substantial hypoechoic regions relating to the overlying rib anatomy. We did not observe any cases in which the modification of pixels affected the visual interpretability of the videos. In sum, these data suggest that this anonymization step is specific to fixed image elements such as ePHI and does not substantially affect data contained within the US cone.

Table 1. Summary Counts for US Samples Used in this Study

Category	n
Manufacturer	
Philips	50
SonoSite	50
Transducer	
Linear	37
Curvilinear	25
Transvaginal	20
Phased array	18
Anatomy	
Renal	20
Pelvic	20
Abscess	20
Lung	20
Cardiac	20

Discussion

The US-Pro software enables efficient local US video preprocessing with robust ePHI removal and accurate dynamic cropping. This program enables data export in video or high-performance machine-learning–compatible formats. We anticipate scenarios in which researchers use this pipeline to anonymize US videos to make medical data accessible in-house or to the larger scientific community[27] without sharing PHI.[41] Although the described algorithms were tested on data from varied organ systems captured from different machines and with all major transducer types, individual US clips should be reviewed by experts to ensure that their quality and content are appropriate for downstream analysis. For example, point-of-care US practice sometimes includes the capture of multiple clips of a single anatomic feature with various overlaid elements (e.g., color Doppler). User input will be required to identify those images applicable to the research question at hand. Furthermore, a key limitation of this approach is that it is only compatible with US videos. The use of US stills in this pipeline will require the integration of a separate image-processing platform.[11]

Distribution within the Docker environment greatly facilitates method adoption and reproducibility within the larger community, since all library requirements are bundled within the distribution software.[42] In addition, Docker provides task scalability, as users are able to deploy more system containers as needs emerge. Through integration with a system such as Singularity,[43] this pipeline can be securely deployed in an academic high-performance computing cluster, enabling manipulation of large data sets within a local system. We anticipate that the US-Pro application will serve as a model for future collaborative scientific computing efforts.

References

1. Moore CL, Copel JA. Point-of-care ultrasonography. *N Engl J Med* 2011; 364:749–757.
2. Hall MK, Taylor RA, Luty S, Allen IE, Moore CL. Impact of point-of-care ultrasonography on ED time to disposition for patients with nontraumatic shock. *Am J Emerg Med* 2016; 34:1022–1030.
3. Via G, Hussain A, Wells M, et al. International evidence-based recommendations for focused cardiac ultrasound. *J Am Soc Echocardiogr* 2014; 27:683.e1–683.e33.
4. Shah VP, Tunik MG, Tsung JW. Prospective evaluation of point-of-care ultrasonography for the diagnosis of pneumonia in children and young adults. *JAMA Pediatr* 2013; 167:119–125.
5. Tayal VS, Hasan N, Norton HJ, Tomaszewski CA. The effect of soft-tissue ultrasound on the management of cellulitis in the emergency department. *Acad Emerg Med* 2006; 13:384–388.
6. Lensing AW, Prandoni P, Brandjes D, et al. Detection of deep-vein thrombosis by real-time B-mode ultrasonography. *N Engl J Med* 1989; 320:342–345.
7. Liu RB, Donroe JH, McNamara RL, Forman HP, Moore CL. The practice and implications of finding fluid during point-of-care ultrasonography: a review. *JAMA Intern Med* 2017; 177: 1818–1825.
8. Lewiss RE, Pearl M, Nomura JA, et al. CORD-AEUS: consensus document for the Emergency Ultrasound Milestone Project. *Acad Emerg Med* 2013; 20:740–745.
9. Strickland E. New “ultrasound on a chip” tool could revolutionize medical imaging. IEEE Spectrum: Technology, Engineering, and Science News website. https://spectrum.ieee.org/the-human-os/biomedical/imaging/new-ultrasound-on-a-chip-tool-could-revolutionize-medical-imaging. Accessed November 28, 2017.
10. Smith BC. ClipDeidentifier. Ultrasound of the Week website; 2015. https://www.ultrasoundoftheweek.com/clipdeidentifier/.
11. Monteiro E, Costa C, Oliveira JL. A de-identification pipeline for ultrasound medical images in DICOM format. *J Med Syst* 2017; 41:89.
12. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521: 436–444.
13. Obermeyer Z, Emanuel EJ. Predicting the future: big data, machine learning, and clinical medicine. *N Engl J Med* 2016; 375: 1216–1219.
14. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; 316:2402–2410.
15. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542:115–118.
16. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv* 2017; 1711.05225.
17. Burlina P, Billings S, Joshi N, Albayda J. Automated diagnosis of myositis from muscle ultrasound: exploring the use of machine learning and deep learning methods. *PLoS One* 2017; 12:e0184059.
18. Han S, Kang HK, Jeong JY, et al. A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Phys Med Biol* 2017; 62:7714–7728.
19. Blessberger H, Binder T. Two dimensional speckle tracking echocardiography: basic principles. *Heart* 2010; 96:716–722.
20. Zhang J, Gajjala S, Agrawal P, et al. A Web-deployed computer vision pipeline for automated determination of cardiac structure and function and detection of disease by two-dimensional echocardiography. *arXiv* 2017; 1706.07342.
21. Karpathy A, Todorici G, Shetty S, et al. Large-scale video classification with convolutional neural networks. In: Proceedings of the 2014 I.E. Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2014:1725–1732.
22. Shin JY, Tajbakhsh N, Hurst RT, Kendall CB, Liang J. Automating carotid intima-media thickness video interpretation with convolutional neural networks. In: Proceedings of the 2016 I.E. Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2016; 2526–2535.
23. Chollet F. Keras. *Keras website*; 2015. https://keras.io/.
24. Bastien F, Lamblin P, Pascanu R, et al. Theano: new features and speed improvements. *arXiv* 2012; 1211.5590.
25. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. *arXiv* 2016; 1603.04467.
26. NVIDIA. GPU cloud computing solutions from NVIDIA. NVIDIA website. https://www.nvidia.com/en-us/data-center/gpu-cloud-computing/. Accessed November 28, 2017.
27. Carpenter J. May the best analyst win. *Science* 2011; 331:698–699.
28. Begley CG, Ioannidis JP. Reproducibility in science. *Circ Res* 2015; 116:116–126.
29. Yaffe MB. Reproducibility in science. *Sci Signal* 2015; 8:eg5.
30. Docker. Docker documentation. Docker website. https://docs.docker.com/. Accessed November 28, 2017.
31. Folarin, A. A., Dobson, R. J. & Newhouse, S. J. NGSeasy: a next generation sequencing pipeline in Docker containers. *F1000 Res* 2015; 7104.1.
32. Django. The Web framework for perfectionists with deadlines. Django website. https://www.djangoproject.com/. Accessed November 29, 2017.
33. Pal NR, Pal SK. A review on image segmentation techniques. *Pattern Recogn* 1993; 26:1277–1294.
34. Simonyan K, Zisserman A. Very Deep convolutional networks for large-scale image recognition. *arXiv* 2014; 1409.1556.

35. Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 2016; 35:1299–1312.
36. Python Software Foundation. Pickle: Python object serialization. Python Software Foundation website. https://docs.python.org/3/library/pickle.html. Accessed January 8, 2018.
37. The HDF Group. HDF home. The HDF Group website. https://www.hdfgroup.org/. Accessed November 29, 2017.
38. Mason CE, Zumbo P, Sanders S, et al. Standardizing the next generation of bioinformatics software development with BioHDF (HDF5). In: *Advances in Computational Biology*. New York, NY: Springer; 2010:693–700.
39. h5py. Datasets. h5py website. http://docs.h5py.org/en/latest/high/dataset.html. Accessed November 29, 2017.
40. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20:37–46.
41. Li P, Xu C, Luo Y, et al. CareNet: building regulation-compliant home-based healthcare services with software-defined infrastructure. In: Proceedings of the 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2017:373–382.
42. Schulz WL, Durant TJS, Siddon AJ, Torres R. Use of application containers and workflows for genomic data analysis. *J Pathol Inform* 2016; 7:53.
43. Kurtzer GM, Sochat V, Bauer MW. Singularity: scientific containers for mobility of compute. *PLoS One* 2017; 12: e0177459.