Extending the compression range of biomedical images for machine vision analysis

Edgar S. Paulo*†, Lucas A. Thomaz*†, Luís M. N. Távora*†, Pedro A. A. Assuncao*† and Sérgio M. M. Faria*†
*Instituto de Telecomunicações, Leiria, Portugal
†ESTG, Polytechnic of Leiria, Leiria, Portugal

***Abstract*—The growing adoption of biomedical machine vision algorithms to perform detection, segmentation, and classification tasks, is driving a shift in compression paradigms, progressively replacing perceptual quality by performance of machine vision tasks as the target encoding optimization. Therefore, improving task performance rather than image quality has become a new research problem in biomedical image compression. This paper presents a contribution to extend the useful compression range from lossless to lossy while keeping the performance of biomedical machine vision algorithms. Automatic detection of mitochondrias in electron microscopy images, using a learning-based network (YOLO), is the case-study investigated in this work. Two types of new results are presented in regard to detection performance. In the first one, it is shown that compression ratios up to 15 can be used, for a maximum of 3% of detection loss. Then in the second one, by using compressed images for training, it is shown that the compression range can be increased up to 135 times, while missing less than 5% of the detections.**

***Index Terms*—Mitochondria detection, YOLO, biomedical image coding, automatic detection, biomedical signal processing**

I. Introduction

Advances in biomedical imaging technology are leading to a significant increase in the amount of generated data. It is estimated that nowadays, healthcare industry accounts for more 30% of the data globally produced [1], and this increase comes accompanied of new challenges, namely the capacity to process, interpret, store, and transmit all the information, keeping in mind their associated costs [2], [3]. Also, the growing number of repetitive systematic jobs increased the involvement of specialists in repetitive tasks, prompting the development of machine vision solutions for automated image analysis [4].

This permanent and growing pressure on existing resources comes from two sides. On one hand, the need to capture and preserve medical information, such as in diagnosis. On the other hand, compliance with legal terms very often require data to be stored for long periods of time, eventually leading a scarcity of computational and storage resources [5]. In this context, biomedical image data compression is necessary to cope with such increasing demand.

Authors' e-mails: {edgar.paulo, lucas.thomaz, luis.tavora, amado, sergio.faria}@co.it.pt. This work was supported by the Fundação para a Ciência e a Tecnologia (FCT), Portugal, Programa Operacional Regional do Centro, and by FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/EEA/50008/2020.

While coding algorithms allow to reduce the amount of data used to represent the biomedical information, the inherent distortion of lossy algorithms poses fidelity problems with impact on the usefulness of coded data. This is the reason for using lossless compression algorithms in many medical applications [6]. However, since lossless compression ratios are far from those obtained from lossy coding schemes, more advanced approaches are required to preserve the relevant biomedical information after lossy coding. This is similar to visually lossless coding [7], but using biomedical task performance as the objective measure.

Visual analysis of biomedical information through manual identification and quantification of anatomic and cellular regions is time consuming and prone to human factors. The small size of relevant elements in high resolution images combined with such human factors tend to produce variable results due to subjective analysis. Also, intra- and inter-observer reliability has been acknowledged as a possible limitation that should be properly accounted for [8]–[10].

A solution that has attracted much attention in recent years consists in delegating the analysis task to machine vision algorithms [11], which minimize the problems of human-resource consumption and intra- and inter-operator reliability, thus ensuring increased overall efficiency. In the biomedical field, algorithms like YOLO [12] and U-Net [13], the latter specially developed for biomedical image segmentation, have been successfully used on the detection of the Malaria parasite [14] and breast cancer [15] as well as cell counting [16], among other applications.

Since higher compression ratios lead to higher coding distortion and consequently lower performance of machine vision tasks, two relevant research questions arise: i) what are the upper compression boundaries that still preserve the relevant information for detection algorithms? and ii) how can such boundaries be extended in order to achieve the same performance with higher compression ratios? This paper presents a contribution for answering these two questions. In particular, automatic detection of cellular structures (mitochondria) in electron microscopy images is considered. An evaluation study of compression ratio versus the YOLO's capability to detect mitochondria in electron microscopy images is first presented, establishing an useful upper boundary for compression. Then a new training methodology for learning-based machine vision algorithms is proposed to extend the image compression range

where performance of such algorithms can be maintained above the same acceptable limit.

The remainder of the paper is organized as follows: Section II presents background information on image compression and object detection, as well as related work. In Section III, the processing pipeline is described and results of the experimental assessment presented, including a newly proposed methodology. The concluding remarks are finally presented in Section VI.

II. Background

In clinical environments, the strict and conservative requirements of the DICOM standard [6] have been driving the use and widespread adoption of traditional image coding technologies, such as JPEG2000 and JPEG-LS. However such encoders are not always the best choice, as their level of compression efficiency was already surpassed by more advanced encoders such as HEVC [17]. In fact, as demonstrated in [18], when used for lossless compression of medical images, the HEVC encoder is able to achieve compression gains up to 43% higher than other reference encoders.

Nevertheless, compression ratios obtained with lossless algorithms are quite limited when compared to their lossy counterparts. Even if not transversely to all applications, in some cases the use of lossy compression for medical data, may result in minimal or no negative impact. For instance, in [19], it was shown that visually lossless coding can be used to compress medical data with no negative impact in medical diagnosis.

In recent years, many of the more repetitive and menial tasks in the analysis of medical data, such as cell counting in blood samples, and organelle identification in microscopy images, have been transferred to machine vision algorithms [20]. These algorithms have the advantage of being time invariant and not suffering from intra- or inter-operator variability [21].

In the recent past, research in medical image analysis has been advancing several application areas such as skin lesion segmentation in dermoscopic images [22], (reporting over 93% accuracy and 90% sensitivity), detection and segmentation of the thyroid gland in ultrasound images [23] (with over 85% correct detection and over 95% correct segmentation of the thyroid gland), real time detection of and segmentation of lung nodules in low dosage CT scans [24] (reaching 0.89 sensitivity and 93% precision) and others e.g., detection of cholelithiasis and gallstones in CT scans [25].

Nowadays, one of the most successful deep learning frameworks for image analysis is the YOLO, namely its versions 3 and 4 [12]. This type of learning network is used for object detection and classification of images of various classes. It also has the added advantage of being fast, presenting good performance for new classes (with moderately few samples) after few epochs of training using transfer learning for large datasets. In medical imaging, specifically for the task of organelles (namely mitochondria) detection and segmentation, the work presented in [26] employs the YOLO v4 with great success.

III. Methods and Experimental Setup

As previously mentioned, this study deals with detection of mitochondrias in compressed electron microscopy images. The YOLO object detection framework is used along with two publicly available datasets [27]: Lucchi++ and Kasthuri++. These datasets are re-annotated versions of data originally made available by Lucchi *et al.* [28] and Kasthuri *et al.* [29], which were verified and adjusted by a group of biologists and neuroscientists. This led to significant changes, and the new versions of the datasets present 20% less mitochondria annotations than the original ground-truth. Additionally, the overall area of the annotations increased by 20% for the Lucci++ and 2% for the Kasthuri++ datasets.

The Lucchi++ dataset is provided in two sets, one for training and one for testing, each one containing 165 images with 1024×768 pixels, in 8-bit grayscale. The first set was further (randomly) subdivided into 125 images for training and 40 for validation. The Kasthuri++ dataset has also 2 sets: one with 85 images with 1463 × 1613 pixels, and another containing 75 images with 1334 × 1553 pixels, all in 8-bit grayscale. Also in this case, the first set was randomly subdivided into 61 images for training and 24 for validation, while the second one was used for test.

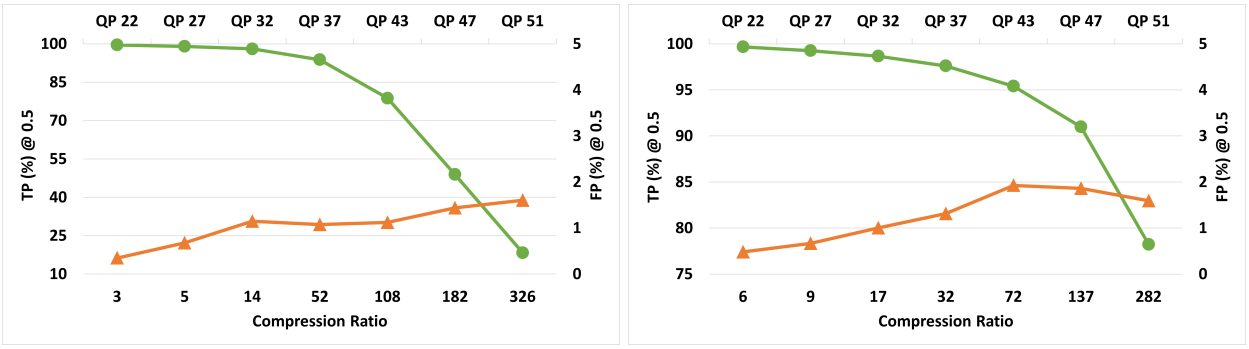
The YOLO network was trained with a maximum of 6000 batches, a batch size of 16, and batch normalisation. The stochastic gradient descent optimizer was used with a momentum term of 0.949 and a learning rate of 1.20×10^{-3}. A step learning rate policy was used, thus after 80% and 90% of the maximum number of batches, the learning rate was decreased by 90%. In the first layer, each image was resized to a resolution of 416×416 pixels.

For data compression, the High Efficient Video Coding (HEVC) standard was used, since it has demonstrated to have a notable performance in medical images compression [18]. Accordingly, the input images were coded using only intra mode, with fixed quantization parameters (QP): 22, 27, 32, 37, 43, 47, and 51, i.e., suppressing possible variations of delta QPs. To evaluate the compression performance, the Compression Ratio (CR) was calculated for each dataset according to

$$\mathrm{CR} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{Size}(\mathrm{ORIG(i)})}{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{Size}(\mathrm{COMP(i)})}, \tag{1}$$

where N is the total number of images in the dataset, $\mathrm{Size}(\mathrm{ORIG(i)})$ is the number of bits of the uncompressed image i, and $\mathrm{Size}(\mathrm{ORIG(i)})$ is the number of bits representing the compressed image.

In this study, objective metrics related to the ability of the YOLO network to detect the mitochondria before and after compression were used. The performance of YOLO in the detection mitochondrias in uncompressed and compressed images was assessed in terms of the following well established indicators: TP (true positives) and FP (false positives); in



(a) Lucci++ dataset.

(b) Kasthuri++ dataset.

Fig. 1. Mitochondria detection performance in compressed images: TP (green circle), and FP (orange triangle).

this type of studies the false negative rate (FN) can be then simply estimated from $FN = 1 - TP$. Using the current implementation of YOLO (v4), the criteria for a successful mitochondria detection (i.e. a TP) is established in terms of the well known IoU indicator, in the case

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \geq 0.5, \tag{2}$$

meaning that the algorithms' output and the annotated reference intersection area should represent at least 50% of their union. In order to avoid spurious "noisy" detections, the network was configured to ignore mitochondrias whose size is less than 50 pixels.

IV. Bounds of Lossy Coding

The aim of the first stage of this study is determine the upper bounds of lossy compression with reduced impact in the YOLO's detection performance. The images in test subsets were encoded using HEVC with different QP values and, for each one, the trained version of YOLO was used to identify the mitochondrias. The results, in terms of average *TP* and *FP* for the different compression ratios, using the prediction over uncompressed images as reference, are presented in Figures 1(a) and 1(b) for Lucchi++ and Kasthuri++, respectively.

The results make it clear that it is indeed possible to compress this type of images with little-to-no impact on the performance of the detection algorithm. In the case of the Lucchi++ dataset, up to a CR of the order of 14 (QP32), the degradation on TP and FP is barely noticeable ($<2\%$); even with CR $\approx$ 52 (QP 37) the average TP was found to be around 95%. The same kind of behaviour was obtained with the Kasthuri++ dataset, where fluctuations $\leq 1\%$ are observed up to CR $\approx$ 17 (QP 32); these results showed a remarkable TP of around 98% for CR $\approx$ 32 (QP 37)[1].

However, if the CR is increased beyond these values, then the detection performance begins to be gradually affected. The algorithm looses some ability to identify the mitochondria region, as indicated by the drop of TP (the increase in FP is only marginal). This behaviour can be understood on the grounds of the coding distortions incurred by compressed images, especially the well known smoothing due to the suppression of higher spatial frequencies. They can be observed in Figure 2, where it is clear that the loss of image detail leads to higher number of undetected mitochondrias. It is important to observe that the algorithm's ability to identify the mitochondrias, although dependent from the overall image quality, is not as sensible as the human vision, as it can be noted by the accurate detection of some mitochondria, even in a blurry image such as Figure 2(b).

For comparison, visual quality metrics related to the human visual system (PSNR and SSIM) obtained from the compression of both datasets with the chosen QPs are shown in Table I. As a rule of thumb, PSNR below 35dB produce visually poor images. In the proposed experiment the detection network was able to detect with very good results even when the PSNR was close to 30dB (QP 32 in Lucci++).

TABLE I
Objective Image Quality per QP

	Lucchi++						
QP	22	27	32	37	43	47	51
CR	3.0	4.7	14.5	52.0	108.1	182.5	325.6
PSNR (dB)	41.2	35.8	30.0	28.2	27.0	25.8	24.4
SSIM	0.996	0.986	0.943	0.913	0.885	0.853	0.800

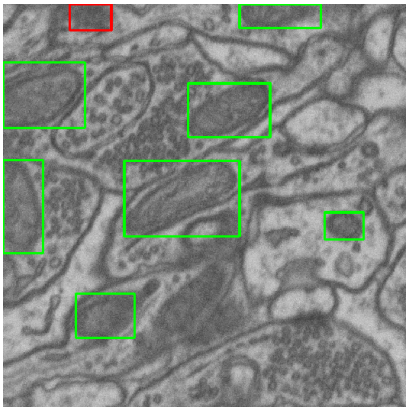
	Kasthuri++						
QP	22	27	32	37	43	47	51
CR	5.8	9.4	17.0	31.8	71.9	137.4	282.1
PSNR (dB)	43.2	38.2	34.2	31.1	27.7	25.6	23.7
SSIM	0.999	0.997	0.993	0.986	0.963	0.904	0.893

In resume, the set of experiments carried out made it clear that it is possible to compress this type of electron microscopy images, up to $\approx 1/15\,th$ of its original size, with only a residual impact in the ability of YOLO to still correctly identify mitochondria regions.

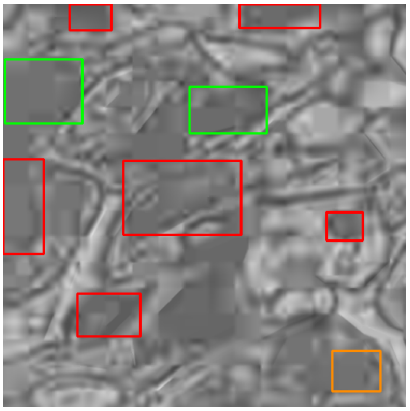
V. Extending the Compression Range

In the light of the results found in section IV, a new approach is devised to extend the useful range of compression ratios by mitigating the effect of higher CRs on the performance of object detection. This is accomplished by

[1]As a reference, the JPEG2000 encoder achieves an average lossless CR of 1.44 and 2.44 on Lucchi++ and Kasthuri++ datasets, respectively.



(a) Image compressed with QP22.



(b) Image compressed with QP51.

Fig. 2. Detail of image 18 of test subset of dataset Lucchi++: TP (Green bounding boxes), FN (red bounding boxes), and FP (orange bounding boxes).

using a different training approach for the learning-based object detector (YOLO) using images coded with various CRs. This is a transfer learning approach plus retraining, which results in fine tuning the object detector by learning how to identify mitochondria in compressed images. In spite of coding distortion, which may remove some relevant features and introduce additional ones, as inferred from Figure 2.

The training procedure used for this new detector follows the same approach as the previous one, but the process starts by loading the weights from the previous stage.The training configurations that differ from the previous YOLO training are the following: learning rate of 5×10^{-5} and maximum number of batches of 10000. These changes in the training parameters were due to the changes in the training set composition. In this new approach each training, validation, and testing set include not only the original images, but also a compressed version of the images initially present in the subset. The following QPs were used for compressing these images: 22, 27, 32, 37, 43, 47, and 51. This means that in this new training, the dataset was expanded eight-fold. Table II shows the number of training images in each dataset.

The results obtained with this new training approach are presented in Figures 3(a) and 3(b) for the Lucchi++ and

TABLE II
NUMBER OF IMAGES IN THE TWO DATASETS

		Training	Validation	Test
Original	Lucchi++	125	40	165
	Kasthuri++	61	24	75
Original + compressed	Lucchi++	1000	320	165
	Kasthuri++	488	192	75

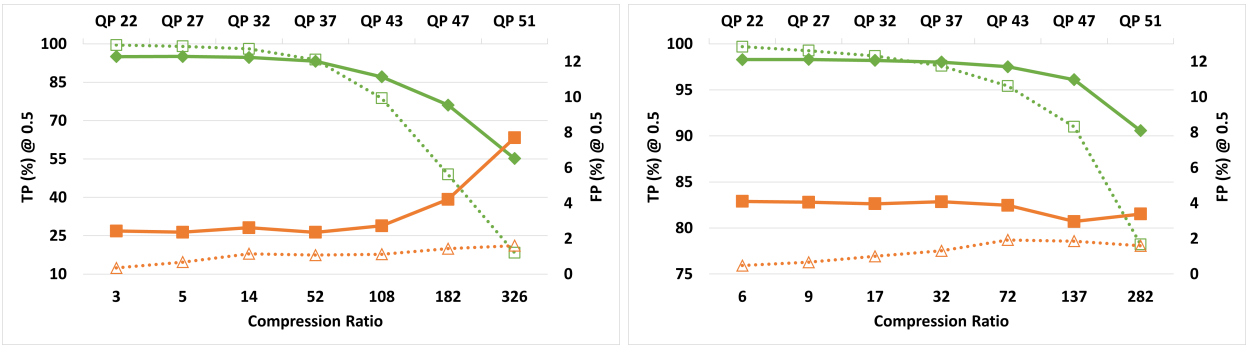
Kasthuri++ datasets, respectively. For comparison purposes, the previous results are also shown, in dashed lines. As can be observed in these figures, with the new proposed training approach it is possible to further extend CRs with little performance degradation; in the Lucchi++ dataset, at CR=108 is still possible to reach about 87% in TP, with FP below 3% (F1_score $\approx$ 92%). Even better figures were obtained for dataset Kasthuri++ where, even for CR=137, TP is around 97% and FP below 4% (F1_score $\approx$ 96%); in terms of quality metrics, according to Table I, this means extending the useful operational range down to around 27dB. It is worth mentioning that, since the network is designed to detect a broader range of objects, its overall performance can be slightly affected for the lower QPs, that is, when the compressed image is still very similar to the original uncompressed one. However, as the level of coding distortion in the images increases, the performance of the proposed approach rapidly surpasses that of the previous one.

In summary, these results demonstrate that automatic mitochondria detection may perform equally well in uncompressed images and compressed up to CRs of 15. This means that lossless CRs can be extended to lossy CR of 15 while still keeping the performance of detection algorithm (YOLO). If the detection network is trained using compressed versions of the dataset, then the compression ratio of 15 can be further extended over 108 in Lucci++ and 135 in Kasthuri++.

VI. CONCLUSION

This paper addresses a current prominent challenge in biomedical imaging applications: how to increase compression ratios of biomedical imaging data, while preserving its essential information for machine vision tasks, such as detection of mitochondrias in electron microscopy images using YOLO. The results demonstrate that it is possible to reduce the amount of data required to represent these type of images, up to $1/15$ of its original size, with barely no impact on the ability to correctly identify the mitochondrias with YOLO. Moreover, with recourse to the proposed training methodology, that uses compressed versions of the original data incorporated in the training set, it is shown that the compression range can be further extended, reaching CRs of the order of 100, with little impact on performance of the detection algorithm.

As future work, the proposed methodology will be used for other biomedical types of images, and also considering the state-of-the-art VVC encoder.



(a) Lucchi++ dataset.

(b) Kasthuri++ dataset.

Fig. 3. Mitochondria detection performance of the proposed method in compressed images: TP (green circle) and FP (orange triangle), proposed method (solid lines), and previous results (dashed lines).

REFERENCES

[1] S. Coughlin, D. Roberts, K. O'Neill, and P. Brooks, "Looking to tomorrow's healthcare today: a participatory health perspective," *Internal Medicine Journal*, vol. 48, pp. 92–96, January 2018.

[2] Alex A. T. Bui, John Darrell Van Horn, et al., "Envisioning the future of 'big data' biomedicine.," *Journal of biomedical informatics*, vol. 69, pp. 115–117, 5 2017.

[3] Sabyasachi Dash, Sushil Kumar Shakyawar, et al., "Big data in healthcare: management, analysis and future prospects," *Journal of Big Data*, vol. 6, pp. 54, 12 2019.

[4] Nikita Orlov, David Eckley, Lior Shamir, and Ilya Goldberg, "Machine vision for classifying biological and biomedical images," *Visualization, Imaging, and Image Processing*, pp. 192–196, 03 2008.

[5] European Commission, Agriculture Consumers, Health, Food Executive Agency, J Hansen, P Wilson, E Verhoeven, M Kroneman, M Kirwan, R Verheij, and E Veen, *Assessment of the EU Member States' rules on health data in the light of GDPR*, Publications Office, 2021.

[6] Medical Imaging & Technology Alliance, "DICOM PS3.10 2022a - Media Storage and File Format for Media Interchange," Tech. Rep., Medical Imaging & Technology Alliance, NEMA, 2022.

[7] L. Prangnell, "Visually lossless coding in hevc: A high bit depth and 4:4:4 capable jnd-based perceptual quantisation technique for hevc," *Signal Processing: Image Communication*, vol. 63, pp. 125–140, 2018.

[8] H. Bahurel-Barrera, G. Assie, et al., "Inter- and intra-observer variability in detection and progression assessment with MRI of microadenoma in Cushing's disease patients followed up after bilateral adrenalectomy," *Pituitary*, vol. 11, no. 3, pp. 263–269, 2008.

[9] Ashleigh Storr, Christos A. Venetis, et al., "Inter-observer and intra-observer agreement between embryologists during selection of a single Day 5 embryo for transfer: a multicenter study," *Human Reproduction*, vol. 32, no. 2, pp. 307–314, 01 2017.

[10] Riitta Salonen, Arto Haapanen, and Jukka T. Salonen, "Measurement of intima-media thickness of common carotid arteries with high-resolution b-mode ultrasonography: Inter- and intra-observer variability," *Ultrasound in Medicine & Biology*, vol. 17, no. 3, pp. 225–230, 1991.

[11] Dinggang Shen, Guorong Wu, and Heung-Il Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248, 6 2017.

[12] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, "Yolov4: Optimal speed and accuracy of object detection," *CoRR*, vol. abs/2004.10934, April 2020.

[13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[14] Fetulhak Abdurahman, Kinde Anlay Fante, and Mohammed Aliy, "Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models," *BMC Bioinformatics*, vol. 22, no. 1, December 2021.

[15] Mohammed A. Al-masni, Mugahed A. Al-antari, et al., "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 85–94, April 2018.

[16] T. Falk, D. Mai, R. Bensch, et al., "U-net – deep learning for cell counting, detection, and morphometry," *Nature Methods*, vol. 16, pp. 67–70, 2019.

[17] Francisco Cunha, Sérgio Faria, João Parracho, Lucas Thomaz, and Pedro Assunção, "Data compression algorithms for biomedical images," in *Imaging Modalities for Biological and Preclinical Research: A Compendium, Volume 2*, 2053-2563, pp. III.4.f–1 to III.4.f–11. IOP Publishing, 2021.

[18] Saurin S. Parikh, Damian Ruiz, Hari Kalva, et al., "High bit-depth medical image compression with hevc," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, pp. 552–560, 3 2018.

[19] D. Wu, D.M. Tan, and H.R. Wu, "Visually lossless adaptive compression of medical images," in *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, 2003, vol. 1, pp. 458–463 Vol.1.

[20] Cheolsoo Park, Clive Cheong Took, and Joon-Kyung Seong, "Machine learning in biomedical engineering," *Biomedical Engineering Letters*, vol. 8, pp. 1–3, 2 2018.

[21] Martijn P.A. Starmans, Sebastian R. van der Voort, Jose M. Castillo Tovar, Jifke F. Veenland, Stefan Klein, and Wiro J. Niessen, "Chapter 18 - radiomics: Data mining using quantitative medical image features," in *Handbook of Medical Image Computing and Computer Assisted Intervention*, S. Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger, Eds., The Elsevier and MICCAI Society Book Series, pp. 429–456. Academic Press, 2020.

[22] Halil Murat Ünver and Enes Ayan, "Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm," *Diagnostics*, vol. 9, pp. 72, 7 2019.

[23] Xiangqiong Wu, Guanghua Tan, Ningbo Zhu, et al., "Cachetrack-yolo: Real-time detection and tracking for thyroid nodules and surrounding tissues in ultrasound videos," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, pp. 3812–3823, 10 2021.

[24] Sindhu Ramachandran S., Jose George, et al., "Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans," in *Medical Imaging 2018: Computer-Aided Diagnosis*, Nicholas Petrick and Kensaku Mori, Eds. International Society for Optics and Photonics, 2018, vol. 10575, pp. 347 – 355, SPIE.

[25] Shanchen Pang, Tong Ding, Sibo Qiao, et al., "A novel yolov3-arch model for identifying cholelithiasis and classifying gallstones on ct images.," *PloS one*, vol. 14, pp. e0217647, 2019.

[26] Asmaa Haja and Lambert R. B. Schomaker, *A Fully Automated End-to-End Process for Fluorescence Microscopy Images of Yeast Cells: From Segmentation to Detection and Classification*, vol. 784 LNEE, Springer Science and Business Media Deutschland GmbH, 4 2022.

[27] Vincent Casser, Kai Kang, et al., "Fast mitochondria detection for connectomics," *Proceedings of Machine Learning Research*, 12 2018.

[28] Aurélien Lucchi, Yunpeng Li, et al., *Structured Image Segmentation Using Kernelized Features*, vol. 7573 LNCS, Springer, Berlin, Heidelberg, 2012.

[29] Narayanan Kasthuri, Kenneth Jeffrey Hayworth, Daniel Raimund Berger, et al., "Saturated reconstruction of a volume of neocortex," *Cell*, vol. 162, pp. 648–661, 7 2015.