DeepFoci: Deep learning-based algorithm for fast automatic analysis of DNA double-strand break ionizing radiation-induced foci

Tomas Vicar [a,b,c], Jaromir Gumulec [c], Radim Kolar [a], Olga Kopecna [b], Eva Pagacova [b], Iva Falkova [b], Martin Falk [b,*]

[a] Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 3058/10, Brno, Czech Republic
[b] Czech Academy of Sciences, Institute of Biophysics, v.v.i, Department of Cell Biology and Radiobiology, Kralovopolska 135, Brno, Czech Republic
[c] Department of Pathological Physiology, Faculty of Medicine, Masaryk University, Kamenice 5, Brno, Czech Republic

ARTICLE INFO

ABSTRACT

DNA double-strand breaks (DSBs), marked by ionizing radiation-induced (repair) foci (IRIFs), are the most serious DNA lesions and are dangerous to human health. IRIF quantification based on confocal microscopy represents the most sensitive and gold-standard method in radiation biodosimetry and allows research on DSB induction and repair at the molecular and single-cell levels. In this study, we introduce DeepFoci – a deep learning-based fully automatic method for IRIF counting and morphometric analysis. DeepFoci is designed to work with 3D multichannel data (trained for 53BP1 and γH2AX) and uses U-Net for nucleus segmentation and IRIF detection, together with maximally stable extremal region-based IRIF segmentation.

The proposed method was trained and tested on challenging datasets consisting of mixtures of nonir-radiated and irradiated cells of different types and IRIF characteristics – permanent cell lines (NHDFs, U-87) and primary cell cultures prepared from tumors and adjacent normal tissues of head and neck cancer patients. The cells were dosed with 0.5–8 Gy γ-rays and fixed at multiple (0–24 h) postirradiation times. Under all circumstances, DeepFoci quantified the number of IRIFs with the highest accuracy among current advanced algorithms. Moreover, while the detection error of DeepFoci remained comparable to the variability between two experienced experts, the software maintained its sensitivity and fidelity across dramatically different IRIF counts per nucleus. In addition, information was extracted on IRIF 3D morphometric features and repair protein colocalization within IRIFs. This approach allowed multiparameter IRIF categorization of single- or multichannel data, thereby refining the analysis of DSB repair processes and classification of patient tumors, with the potential to identify specific cell subclones.

The developed software improves IRIF quantification for various practical applications (radiotherapy monitoring, biodosimetry, etc.) and opens the door to advanced DSB focus analysis and, in turn, a better understanding of (radiation-induced) DNA damage and repair.

1. Introduction

The ability to precisely and rapidly monitor DNA double-strand break (DSB) induction and repair underpins numerous fields of biological, medical and space research (reviewed, e.g., in [1–4]). DNA DSBs are regularly introduced into DNA molecules by ionizing radiation, radiomimetic chemicals and vital cell processes [5–7]. With both DNA strands simultaneously cut, DSBs represent the most serious type of DNA lesions [8], the accumulation of which, if left unrepaired, promotes aging [9], neurodegeneration [10], infertility [11], and other health consequences. Imprecise DSB repair then

causes mutagenesis and may lead to cancer [7,12]. DSB damage induction and repair monitoring thus opens the doors to personalized medicine [3,4,13–15], for instance, coupled with direct radiotherapy effect monitoring [13,16] and rational development of new DNA-damaging [17–20] or DNA-protecting drugs [21–24] needed in medicine [13,16,25] and radiation protection [26,27]. Precise and automated DSB damage monitoring is also valuable for biodosimetry [27–31], for instance, in situations related to mass radiation accidents (terrorist attacks with radioactive materials) or space exploration where astronauts would be exposed to a mixed field of different radiation types [32–38]. Although biodosimetry based on immunofluorescence detection of ionizing radiation-induced (repair) foci (IRIFs) does not replace the methods based on other biomarkers (e.g., chromosomal aberrations), it is currently the most sensitive approach to detect DSB damage, as detailed below.

A revolution in DSB detection has arrived with the discovery of IRIFs that rapidly form at DSB sites after damage, which currently serve as their most sensitive markers (reviewed, e.g., in [2–4]). One of the early events at DSB sites is the phosphorylation of histone H2AX at serine 139 (referred to as γH2AX), which eventually spreads over 2 Mb of damaged chromatin and leads to the formation of so-called γH2AX foci [39]. γH2AX foci then serve as signaling and structural platforms [40] that attract, in a spatially and temporally organized manner, additional repair proteins to DSB sites. Consequently, IRIFs of different repair proteins, characterized by specific parameters and behavior, can be visualized in cell nuclei by immunofluorescence microscopy. Since the number of IRIFs tightly corresponds with the number of DSBs in most DNA damage situations [41–43], the IRIFs formed by γH2AX or colocalized repair proteins (such as 53BP1) can be considered quantitative DSB markers with a single lesion sensitivity [44]. Although IRIFs are only temporary structures, the advantage of the method is that blood samples obtained from potential victims can be simply fixed at the site of the accident and processed later, even after a long period, in the laboratory.

While flow cytometry offers fast, automated quantification of integrated values of these repair signals in high cell numbers [45], microscopy allows detection of individual IRIFs in the context of their natural chromatin environment in individual cells and analysis of their property development over time [35,46,47]. Characterization of morphological and behavioral parameters of IRIFs formed by repair proteins participating in different DSB repair pathways [48]—such as 53BP1 and γH2AX, which were used in the present manuscript for illustration—opens the doors to exploring spatiotemporal interactions between repair proteins at individual DSB sites, deepening our insights into mechanisms of DSB induction and (mis)repair [13–15,37,49–56].

IRIFs are, by nature, highly dynamic structures. Their number per nucleus and other parameters, such as size, intensity, shape, and border sharpness, change dramatically during the postdamage period as repair proceeds [37,57–60]. Moreover, cell types or even individual cells of the same population show extreme differences in generated DSB/IRIF numbers, IRIF properties, and intensity of the background signal [36,37,60,61]. This fact is attributed to the randomness of damage induction, generation of DSBs that are repaired with unequal efficiencies, heterogeneous cell states, asynchronous repair of individual DSBs, and cell biological variability. In addition, IRIF parameters are influenced by the sample preparation [31,62]. Confocal microscopy image data thus cannot be considered quantitative without careful calibration and detailed knowledge on the exact experimental and biological behavior of the cell type studied. Simplistic intensity thresholding or approaches based on predefined parameters thus frequently fail, making the correct identification of IRIFs impossible. Accordingly, *The Second γH2AX-Assay Inter-Comparison Exercise* carried out in

the framework of the European Biodosimetry Network (RENEB) [30], other available literature sources, and our experience have demonstrated that manual inspection of images by an expert eye still ensures more precise IRIF identification than automatic software algorithms. Nevertheless, visual quantification of IRIFs is extremely time-consuming and difficult even for a trained eye. Moreover, unless all data are analyzed by a single observer, which is practically impossible, the results may suffer from dramatic variations [28]. Hence, the results obtained by different observers and/or laboratories can only be compared with extreme caution [31,63,64]. This unsatisfactory situation means that without suitable software, the evaluation of large image datasets, as generated, for instance, in the case of mass radiation accidents, remains unrealistic. This problem strongly complicates other practical (e.g., medical) applications and research.

Moreover, the information on architectural IRIF properties, such as the focus size, intensity, and shape, is left unexplored by visual-only evaluation. The IRIF architecture has been recognized as an important factor regulating DSB repair processes and potentially participates in decision-making for a particular repair mechanism (pathway) at individual DSB sites [37,38,56,65–67]. Architectural IRIF defects or their enhanced presence often appear in cells affected by cancer [36,37,68,69], precancerous syndromes [14,15,69] and aging [70]. Hence, the improved ability of automatic software detection coupled with detailed characterization of IRIFs formed by individual repair proteins would be immediately recognized in numerous research fields as well as important practical areas of human activity related to DNA damage (e.g., medicine, radiation protection and space exploration).

Several strategies to segment IRIFs (usually γH2AX or 53BP1) have recently been published [71–75]. Focinator [71,76], FindFoci [72], the method proposed by Feng et al. [73], Foco [74], AutoFoci [75] and FocAn [76] represent the most important open-source attempts. Commercial software packages developed for microscopy image processing by microscope-providing companies are not focused on IRIFs specifically, and the ongoing effort to develop new open-source IRIF analysis platforms clearly demonstrates that many important issues have not been solved satisfactorily.

To conclude, while automatic IRIF quantification (and nucleus segmentation) with only simple processing techniques, such as thresholding, is inefficient due to variability of the fluorescence intensity and other IRIF parameters between the cells and experiments, massive quantities of image data related to emergency biodosimetric events or required to meet the research requirements preclude manual analysis of IRIFs. Even if there are sufficient human resources for this purpose, the results of different evaluators (even from the same laboratory) usually suffer from a strong subjective bias and are therefore hardly comparable. Moreover, IRIF parameters cannot be quantified, except for the number.

In the present manuscript, we introduce DeepFoci, a novel robust software based on machine (deep)-learning strategies for fully automated identification and characterization of IRIFs formed by different repair proteins. The software overcomes serious shortcomings associated with IRIF detection as described above and allows segmentation of cell nuclei and IRIFs with high fidelity, even in the case of challenging cell specimens of dramatically different quality as they appear in daily practice. The precision, specificity and reproducibility of the procedures are further enhanced by dual DSB labeling and colocalization analysis of two selected independent DSB markers [21,35]. At the same time, this approach allows studies on spatiotemporal interactions between IRIF proteins during repair. The software was successfully trained and tested on extremely challenging datasets based on tumor cell primary cultures; in all cases, DeepFoci exhibited performance comparable to that of a careful, time-demanding manual analysis by an experienced expert.

2. Methods

2.1. Cells and cell culturing

The following cells were used: (1) Normal and cancerous standard permanent cell lines represented by primary normal human dermal fibroblasts (NHDFs, PromoCell, Heidelberg, Germany) isolated from the dermis of the juvenile foreskin or adult skin and highly radioresistant U-87 glioblastoma cells (ATCC HTB-14, LGC Standards, United Kingdom), respectively; and (2) Tumor and tumor-adjacent primary cell cultures isolated from patients with squamous cell cancer of head and neck (tumor and tumor-adjacent tissues). While the permanent cell lines represented biologically homogeneous and technically relatively easy samples, tumor cell primary cultures were considered highly heterogeneous and challenging samples.

The protocol for patient primary culture isolation was described in [78]. The primary culture was cultivated in RPMI-1640 medium with Pen/Strep antibiotic solution (PAA Laboratories GmbH, Austria) and 10% fetal bovine serum (FBS, Biochrom, USA), at 37 °C and 5.0% CO_2 in a humidified atmosphere up to 50% confluence. NHDF and U-87 cells were grown in Dulbecco's modified essential medium (DMEM, Life Technologies) supplemented with 10% fetal calf serum (FCS) and a 1% gentamicin–glutamine solution (all reagents from Sigma-Aldrich).

2.2. Irradiation

The cells were irradiated with different doses of γ-rays produced by a ^{60}Co irradiator (Chisostat, Chirana, CR) at the Institute of Biophysics of the Czech Academy of Sciences, Brno, Czech Republic. Permanent cell cultures – NHDFs and cancer U-87 cells – were irradiated with increasing single doses of 0.5, 1, 2, 4 or 8 Gy (D = 1 Gy/min) γ-rays. Patient-derived primary cultures were irradiated with a single dose of 2 Gy due to the different characteristics of the intended experiments and lack of material, using the same irradiation conditions as those of the permanent cell lines. The cells were irradiated in culture flasks containing a normal atmosphere and appropriate culturing medium and placed in a thermostable box to keep the temperature of the cells close to 37 °C. Irradiated cells were immediately returned to the thermostat (37 °C, in a normal atmosphere), and DSBs were repaired for the indicated periods of time postirradiation (PI). Subsequently, the cells were detached from the culturing flask bottoms, attached to glass microscopy slides for 5 min without dehydration, spatially (3D) fixed, immunolabeled, and visualized by confocal microscopy as described in the particular paragraphs below. The irradiation procedure and sample processing were set to resemble real accident conditions.

2.3. Cell fixation and immunostaining

Aliquots of nonirradiated cells (0 min PI) and irradiated cell samples were washed in phosphate-buffered saline (PBS) and spatially (3D) fixed with 4% buffered paraformaldehyde for 10 min at room temperature (RT) at different periods of time PI—30 min, 8 h and 24 h PI. Subsequently, cells were permeabilized with 0.2% Triton X-100/PBS for 15 min and immunolabeled for IRIFs. Two combinations of primary antibodies were used for immunofluorescence detection: anti-phospho-histone H2AX (mouse, clone JBW301; Merck Millipore, Darmstadt, Germany, cat. no.: 05–636; 1:400) + anti-53BP1 (rabbit; Cell Signaling Technology, Danvers, MA, USA, cat. no.: 4937; 1:400) or anti-phospho-histone H2AX (mouse; Merck Millipore; 1:400).

Among other IRIFs and DSB markers, γH2AX foci and 53BP1 foci were selected for the following purposes: γH2AX foci have been used as DSB markers in numerous studies and can indicate changes in chromatin structure that appear at DSB sites during DSB repair. 53BP1 protein participates in DSB repair and the decision-making process for nonhomologous end joining (NHEJ) or homologous recombination (HR) at particular DSB sites. Since 53BP foci colocalize well with γH2AX foci and are formed in a kinetically similar manner to these foci, 53BP1 and γH2AX foci are often codetected to enhance the fidelity and reliability of DSB quantification [21,35].

The immunodetection procedure used was described previously [21,35]. Briefly, after incubation with primary antibodies (overnight at 4 °C), a mixture of secondary antibodies was applied for 1 h (RT). Primary antibodies were visualized by the mixture of FITC-conjugated donkey anti-mouse and Cy3-conjugated donkey anti-rabbit (both Jackson ImmunoResearch Laboratories, West Grove, PA, cat. nos.: 715–095-150 and 711–165-152) applied in 1:100 and 1:200 dilutions, respectively (30 min incubation at RT in the dark). Alternatively, anti-mouse Alexa Fluor 647 and anti-rabbit Alexa Fluor 568 (Thermo Fisher Scientific) secondary antibodies were used, which are directly compatible with both confocal microscopy and single-molecule localization microscopy (SMLM). The antibodies were diluted in sterile donkey serum (1:400 and 1:200, respectively; cat. no.: P30-0101, Pan Biotech GmBH) and applied to the cells for 30 min (RT, in the dark). After incubation, the cells were washed three times in 1 × PBS for 5 min.

The cell nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, 5 min at RT) provided as Duolink *In Situ* Mounting Medium with DAPI (DUO82040; Sigma-Aldrich; now Merck, Darmstadt, Germany) and diluted to a concentration of 1:20,000 [79]. Afterward, the slides with cells were washed three times in 1 × PBS for 5 min each. Finally, the coverslips were air-dried, and the cells were embedded in ProLong Gold (Thermo Fisher Scientific). ProLong Gold was left to polymerize for 24 h in the dark at RT. After complete polymerization, the slides were sealed with nail polish and stored in the dark at 4 °C. Alternatively, nuclear chromatin was counterstained with 1 µM TO-PRO-3 (Molecular Probes, Eugene, OR) in 2 × saline sodium citrate (SSC), prepared fresh from a stock solution. After brief washing in 2 × SSC, Vectashield medium (Vector Laboratories, Burlington, Canada) was used for final mounting of the samples.

2.4. Confocal microscopy

A Leica DM RXA microscope equipped with a DMSTC motorized stage, piezo z-movement, MicroMax CCD camera, CSU-10 confocal unit and 488, 562, and 714 nm laser diodes with acousto-optic tunable filters (AOTFs) was used to acquire detailed cell images with a 100 × oil immersion Plan Fluotar lens (NA 1.3) with a Z step size of 0.3 µm; usually, fifty serial optical sections were captured along the z-axis. The equipment was controlled by the Acquiarium software developed by [80]. The resulting images were 90.0 × 67.2 × 15 µm xyz (1392 × 1040 × 50 px). The exposure times for individual channels were experimentally set to suitable values that ensured comparable imaging, even for samples with different labeling efficiencies. All samples were maintained at a constant temperature of 20 °C to reduce microscope overheating and thus prevent an increase in chromatic aberrations (shifts among the R, G and B channels).

2.5. Datasets for software analyses

Two datasets were used in the study. Dataset (1) was based on patient-derived primary cell cultures prepared from squamous cell cancer and morphologically normal tissues adjacent to the tumor taken from patients suffering from head and neck cancer. The data-

set consisted of several cell types: a) tumor cells, b) tumor-associated fibroblasts, and c) cells from morphologically normal tissues. All cell types were fixed at different periods of time (0 (nonirradiated control), 0.5, 8 or 24 h PI) after exposure to 2 Gy γ-rays. Overall, this dataset consists of 389 samples (containing tumor cells, tumor-associated fibroblasts, or normal cells analyzed at different time points after exposure to different radiation doses) from 77 patients, where each sample contains ∼32 fields of view (FOVs), with 12,650 FOVs in total. The dataset was divided into two subsets: (1.1) for training, validation and testing the nucleus segmentation (237/10/30 FOVs; 1090/89/139 cells) and (1.2) for training, validation and testing the colocalized IRIF segmentation (239/60/100 FOVs; 1047/243/343 cells). The representation of cells in the two subsets with respect to the cell type, patient, radiation dose, and PI time (i.e., DSB repair duration) was randomized to make the datasets as diverse as possible in order to train the most universal method and ensure the most reliable evaluation of real, often very heterogenic cell samples. Validation data were used for evaluations during model training and for adjustment of training parameters. Moreover, postprocessing parameters were set optimally for the validation data. The testing dataset was used to evaluate the proposed method.

Dataset (2) was composed of multiple types of differently treated cell lines to establish a highly challenging dataset maximally reflecting high biological and technical variability between samples, as may appear in research or clinical practice. The dataset contained a) mesenchymal NHDFs coming from a standard permanent cell line (331 annotated FOVs) and b) radioresistant U-87 glioblastoma cells coming from a standard permanent cell line (348 annotated FOVs). NHDF and U-87 cells received 0.5, 1, 2, 4 or 8 Gy γ-rays and were fixed at 0.5 h and 8 h PI. Dataset (2) was used to train single-channel detection of γH2AX and 53BP1 channels in multichannel images and to assess the robustness of segmentation procedures.

2.6. Ground truth generation

Manual annotation of nuclei and IRIFs is required for convolutional neural network (CNN) training and performance evaluation. As both nucleus segmentation and IRIF detection are performed in 3D, manual labeling with available labeling tools is problematic. For this reason, a customized labeling graphical user interface (GUI) was created in MATLAB, enabling easier 3D data labeling. The tool for nucleus segmentation is based on the presegmentation of an image with the simple linear iterative clustering (SLIC) superpixel approach [81], where superpixels are then labeled in the GUI by the user. For detection training, IRIFs were predetected with the same algorithm as that for the final segmentation, where the CNN was replaced by a simple local maxima detector applied on the colocalization image. This detector was set to high sensitivity to capture all potential IRIFs. At this point, the user selects real foci from IRIF proposals in the 2D projection image, where its 3D coordinates are taken from the detector. Nuclei and IRIFs for training were manually annotated in 3D by one expert. IRIFs for testing of the algorithm were labeled in 2D projection.

2.7. Evaluation metrics

To evaluate the accuracy of nucleus segmentation in 3D, the SEG score (objectwise intersection over union (IoU)) was used [81]. To calculate the SEG score, the IoU (also known as the Jaccard index) is needed, which is defined in equation (1):

$$IoU = \frac{X \cap Y}{X \cup Y} \tag{1}$$

where X and Y are manually annotated segmentation masks and predicted segmentation masks, respectively. For every manually annotated object, the segmented object with the largest IoU is found. Next, the average IoU of all manually annotated objects is calculated. If the IoU for any manually annotated object is smaller than 0.5, then the IoU for this object is set to 0. This approach ensures that each manually annotated object can be paired with only one segmented object. The resulting SEG is an average of the IoUs for all manually annotated objects. To evaluate the detection accuracy of individual IRIFs, the Dice coefficient (F1-score) was used, as defined as equation (2):

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{2}$$

where TP is the number of true positive IRIFs, FP is the number of false positives and FN is the number of false negatives.

2.8. Ethical declarations

The study was conducted in accordance with the 1964 Declaration of Helsinki and all subsequent revisions thereof. Moreover, this study was approved by the ethical committee of St. Anne's Faculty Hospital, Brno.

3. Results

To overcome important shortcomings of currently available procedures for DNA repair focus (IRIF) analysis, we have developed DeepFoci, a novel software tool based on artificial neural networks and deep learning that allows fully automated detection, quantification and analysis of these structures in the context of their natural environment, i.e., within the architecture of the cell nucleus (chromatin). DeepFoci is written in MATLAB and Python and is primarily focused on precise 3D segmentation of IRIFs and cell nuclei in large datasets of confocal microscopy images and consequent analysis of recorded data. To fully benefit from the software abilities, IRIF visualization with fluorescently tagged antibodies against two different DSB markers was applied. This dual labeling approach improves the precision of DSB quantification and, at the same time, allows studying the spatiotemporal relationships among γH2AX (or other epigenetic modifications), repair proteins of interest, and chromatin architecture/function at individual DSB sites. Nevertheless, analysis based on staining with a single IRIF marker (e.g., γH2AX) is also possible if preferred based on the characteristics of an experiment or practical situation. In the present manuscript, γH2AX and 53BP1 were selected as IRIF markers—γH2AX because of its widespread usage for this purpose and 53BP1 protein due to its involvement in both main DSB repair pathways (NHEJ and HR) [83,84]. Moreover, γH2AX and 53BP1 foci differ in their morphological features but share similar formation-decomposition kinetics and extensively colocalize with each other. The protein RAD51 is a DSB repair protein selectively involved in HR and, as such, colocalizes with only a subset of γH2AX foci. Foci of RAD51 appear at a later PI time than γH2AX and 53BP1, and their morphology leads to more difficultly in software analysis.

Images were preprocessed with a 5 × 5 × 1 median filter, Gaussian filter with sigma 1. The proposed IRIF analysis method consists of three main steps (see Fig. 1): (1) initial instance segmentation of single nuclei with a CNN, (2) detection of individual IRIFs again with the CNN and (3) segmentation of detected foci with the maximally stable extremal region (MSER) algorithm.
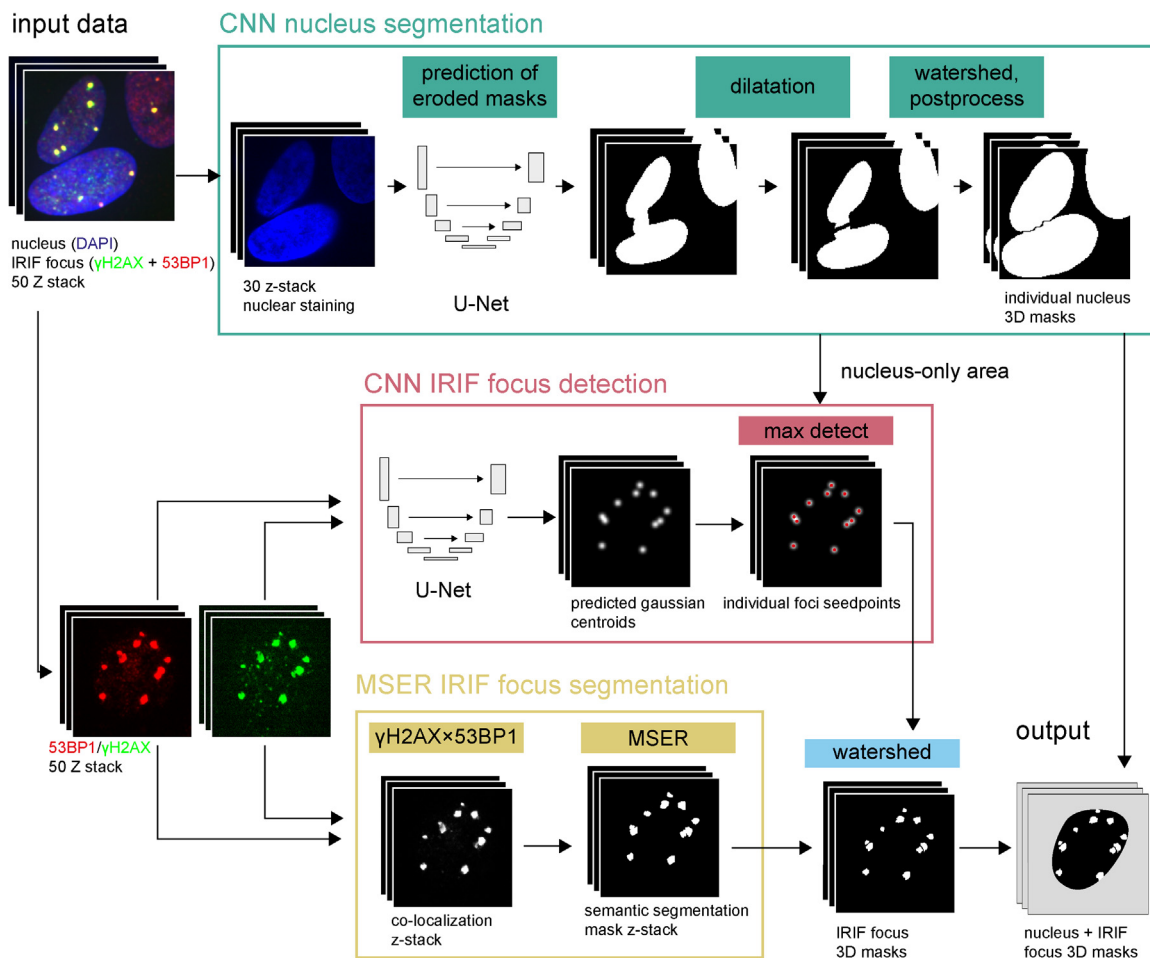
Fig. 1. Block diagram of IRIF detection. Three-channel images are used for a network input: one for nuclear staining and two for IRIF staining, as exemplified by DAPI staining of the nuclei (blue) and immunodetection of γH2AX (green) and 53BP1 (red) IRIFs. The process is divided into three steps. First, 3D nucleus masks are created using a U-Net convolutional neural network (CNN) from the channel for nuclear staining. Second, individual IRIFs are detected with the CNN. Third, individual foci are segmented from a multiplied z-stack composed of the two channels for IRIFs by utilizing a maximally stable extremal region (MSER) detector. The output of these three steps is finally merged into nucleus/IRIF 3D masks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1. Nucleus segmentation

The image-to-image encoder-decoder CNN with the U-Net topology [85] (See Supplementary Fig. 1) was proved to be very powerful for biomedical image segmentation. However, it produces just foreground-background (semantic) segmentation in most standard cases. The standard U-Net architecture does not ensure single nucleus separation, as every error in the boundary voxels results in the connection of neighboring nuclei into one segmented object. Successful separation of individual nuclei was achieved by modifying the network to predict the eroded binary masks. However, this approach can still lead to incomplete separation of individual nuclei due to prediction errors on the boundary between nuclei. Accordingly, 3D CNN prediction is followed by distance transform (DT) and watershed segmentation (applied to the negative DT image) to separate the touching nuclei [86]. Moreover, the DT image is processed with the h-maxima transform and grayscale dilation, which prevents oversegmentation. This step removes the maxima that are close to each other and separated by an insufficient decrease in the image intensity. The minimal distance between the maxima is controlled by the radius of the structuring element and the minimal image intensity decrease defined by the h parameter of the h-maxima transform. Afterward, the resulting image is dilated to compensate for the initial erosion of ground truth masks. To prevent nucleus merging, dilation is performed

sequentially for individual nuclei. Both the h parameter and minimal distance between the maxima were optimized with grid search on a validation set.

3.2. IRIF detection

Similarly, 3D U-Net with the same architecture was applied for the detection of individual IRIFs. In this case, an image with the 3D Gaussian function overlaid over the position of each IRIF centroid represented the ground truth for CNN training. Thus, the 3D CNN estimates possible foci in the form of a Gaussian function, which is further postprocessed. Individual foci were detected using maxima detection—local maxima with a value above the threshold were considered detected foci, and the h-maxima transform was utilized to prevent multiple detections of the same IRIFs due to inaccurate CNN prediction. Moreover, grayscale dilation was applied to set the minimal distance between detected IRIFs. The threshold, h parameter and minimal IRIF distance were optimized with grid search on a validation set.

3.3. IRIF segmentation

The MSER algorithm [87] is a segmentation technique that is generally very robust to illumination changes and, therefore, suitable for the segmentation of fluorescence microscopy images of

varying intensity. Extremal regions of an image are defined as the connected components of a thresholded image. MSER produces stable extremal regions of the image, which are stable in the sense of the volume variation w.r.t. changes in the threshold. The minimal allowed stability of the extracted region can be set with two parameters—the threshold step and the maximal relative volume change associated with this step.

This IRIF segmentation approach was applied to the colocalization image created as a multiplication of the two IRIF channels (as represented by γH2AX and 53BP1 signals here). Using MSER, multiple segmentation variants of increasing size for every IRIF were generated. The size was restricted to the maximal IRIF volume. Of the segmentation variants, the largest was selected as a final segmentation mask. The IRIF segmentation produced by MSER was then combined with U-Net IRIF detection employing the seeded watershed transform. The seeded watershed transform [88] was then applied to the colocalization images, where the outputs of IRIF detection served as the seeds and those of MSER served as the foreground mask.

3.4. Implementation details

MATLAB R2021a with Image processing and Deep Learning Toolboxes and VLFeat library (for 3D MSER) [89] was used. The 3D U-Net network [85] with 16 filters in the first layer was employed for both the nucleus segmentation and IRIF detection settings. The whole network had a depth of 4 (with 4 encoder and decoder blocks separated by max-pooling/transposed convolution), with the number of filters doubled for blocks of lower resolution. Each block consisted of 2 layers of convolution, batch normalization and rectified linear unit (ReLU) nonlinearity; the whole architecture is shown in Supplementary Fig. 1. For nucleus and IRIF detection, only the loss functions were different: the Dice loss [90] for nucleus segmentation and the mean-squared error loss for IRIF detection. For higher computational efficiency and the graphics processing unit (GPU) memory limit of nucleus segmentation and IRIF detection, the image volumes were downscaled in the x-y dimensions by a factor of two and in the z dimension from size 50 to size 48 (to size $505 \times 681 \times 48$ px). Images were preprocessed with normalization of each channel based on image upper and lower percentile values; specifically, values from the 0.01–99.99 percentile range were projected to range -0.5–0.5. Augmentation with a selection of random patches in the x-y direction ($96 \times 96 \times 48$), random flips in the x-y direction, random rotation by multiples of 90°in the x-y direction, and multiplication of image pixels by a random value between 0.5 and 1.5 was used.

3.5. Nucleus segmentation evaluation

The SEG instance segmentation measure [82] was adapted for the segmentation of nuclei. Two nuclei were considered matching if the IoU was equal to or greater than 0.5. Each ground truth nucleus was included exactly once to prevent assignment to multiple nuclei. Nucleus segmentation was tested on a dataset consisting of 30 FOVs annotated by a single expert, using the same tool employed to generate the training data. The proposed method achieved a median SEG score of 0.82 (median over FOVs). The representative FOV with a SEG score of 0.80 and the distribution of SEG values (a histogram for all FOVs) are shown in Fig. 3a and b, respectively.

3.6. Colocalized IRIF detection evaluation

The accuracy of automated IRIF detection was compared to that of manual annotation performed by two experienced experts on the maximum-projection images. The Dice coefficient served as the IRIF detection accuracy metric (see Methods). The IRIFs detected by the proposed software procedures and annotated manually were considered mutually matching if their centroids were closer than 1.95 μm (30 px), a value corresponding to maximal dispersion of manual annotations between experts. The IRIF detection performance using DeepFoci is presented in Fig. 2c-e. Manual annotation by two experts enabled not only evaluating IRIF detection by DeepFoci itself but also estimating the minimum difference that can be observed between experts. The difference between manual expert annotations provided a median Dice coefficient of 0.75. The discrepancies between the automated IRIF detection/segmentation by DeepFoci and manual annotations by either of the experts were close to the variability between experts, with a Dice coefficient of 0.64 for expert 1 and 0.70 for expert 2 (Fig. 2f). An example of automated IRIF detection and segmentation is presented in Fig. 2c-e.

3.7. Colocalized IRIF detection method verification and practical applications

The DeepFoci performance was compared with that of two recently published tools for IRIF counting, namely, FocAn [76] and AutoFoci [75], and with that of CellProfiler (universal particle analysis software). For AutoFoci and FocAn, only the IRIF segmentation part (without nucleus segmentation) was applied. For 2D methods (AutoFoci and CellProfiler), the maximum intensity projection images were used. CellProfiler and FocAn were designed for single IRIF fluorescence staining; thus, the procedures were applied to the colocalization images (obtained by multiplication of the two IRIF channels) because this variant achieved the best results. The parameters of these methods were optimized using grid search. For AutoFoci, the most reliable values were searched and set up for *the object evaluation parameter threshold*, *minimal focus distance* and *top-hat structuring element radius*. For FocAn, *the minimal focus distance*, *threshold value* and *neighborhood size of the adaptive threshold* were optimized. CellProfiler employed a simple pipeline (based on [72]) with the *EnhanceOrSuppressFeatures–Enhance Speckles* option and with the *IdentifyPrimaryObjects–Otsu method* and distance local maxima suppression, where *the minimal focus distance* and *Enhance Speckles* parameters were tuned.

Based on the performance on our challenging testing dataset, composed of heterogeneous primary cell cultures derived from squamous cell head and neck cancer patients' tumors (see Methods for details), the highest accuracy among all compared software detection methods was achieved with DeepFoci. The median Dice coefficient for FocAn, AutoFoci, CellProfiler and DeepFoci was 0.22, 0.38, 0.49 and 0.67, respectively (Fig. 3a).

To further compare DeepFoci with these already available software approaches, the correlation between the manually annotated and automatically detected IRIFs was evaluated for nuclei with various IRIF counts, ranging from 0 or only a few in nonirradiated controls to several dozens in cells fixed at 1 h PI. To cover all stages of DSB repair that differ dramatically in the number of IRIFs per nucleus and IRIF parameters, primary tumor cultures were fixed at different periods of time until 24 h PI. Specifically, the fixation times were selected to test the software ability to quantify a) large quantities of morphologically variable IRIFs at the time of their maximum appearance in nuclei (0.5 h PI), b) medium quantities of large but differently diffused late IRIFs (8 h PI) and c) low quantities of only a few persistent IRIFs ("irreparable" DSBs) in cells that almost accomplished repair (24 h PI). Nonirradiated cells (0 min PI) served as the negative controls with no or only a few naturally occurring IRIFs. Among all software approaches (Fig. 3b), the IRIF numbers detected by DeepFoci best correlated (had the most linear dependence) with those obtained for the same corresponding
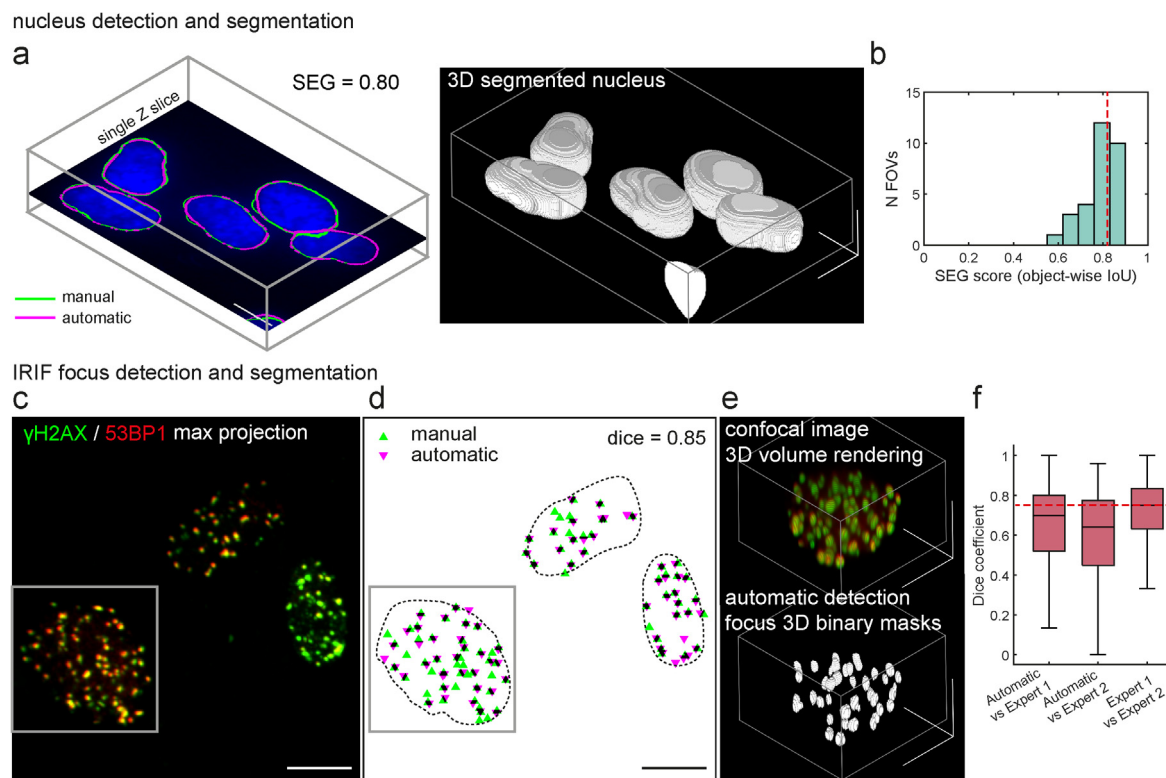
Fig. 2. The results of nucleus segmentation and colocalized IRIF detection. a-b Nucleus segmentation performance; **a**. Comparison of automatic and manual nucleus segmentation. Segmentation results for a single field of view (FOV), single Z slice, NHDF cells, and DAPI staining are shown (left), together with 3D reconstruction (right). **b**. Histogram of the SEG score for nucleus segmentation with 30 segmented FOVs used for testing. The red line indicates the median SEG of all FOVs. **c-f**. IRIF segmentation performance. **c**. IRIFs after 2 Gy γ-ray exposure, 8 h postirradiation, oropharyngeal squamous cancer cells from patients, γH2AX/53BP1 staining, max projection, 100 × magnification. **d** Comparison of manual annotation and DeepFoci detection results. **e**. Top – 3D confocal data of the detail indicated by a gray square in 2c; bottom - binary masks detected by the proposed CNN. **f**. IRIF detection performance, comparison of the automatic result with two manual annotations by experts shown as the median, IQR and min–max in 1.5 IQR. The red line indicates the median Dice coefficient of IRIF detection between two experts. The scale bar in all FOVs indicates 10 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

nuclei by manual expert analysis (average values for individual experts are plotted in Fig. 3b). Moreover, unlike other tested methods, DeepFoci retained both its sensitivity and fidelity over the whole scale of possible IRIF quantities. The detection performance of those methods for representative images with various numbers of IRIFs is shown in Supplementary Fig. 2.

To demonstrate the practical usability of DeepFoci and further test its performance, DSB repair kinetics and IRIF morphology were compared for head and neck squamous cell cancer primary cultures and tumor-adjacent cultures using additional FOVs that were not involved in training. For all the tumor cell primary cultures and tumor-adjacent nontumorous primary cultures, the IRIF numbers peaked at 0.5 h PI, followed by a significant drop at 8 h PI and persistence of only a few IRIFs/nuclei at 24 h PI (Fig. 3c). Such a repair profile (repair kinetics) corresponds well with the profile expected for cells exposed to 2 Gy gamma radiation [13], i.e., the conditions used in the present work.

In addition to the number of IRIFs per nucleus, analyzed as the only parameter in most studies, a wide spectrum of other parameters can be extracted by DeepFoci, including the focus intensity in two color channels, the intensity of chromatin staining at the IRIF site and extent/characteristics of colocalization in all color channels. Furthermore, it was possible to measure the 3D morphometric features of IRIFs and nuclei, such as their volume, solidity, and circularity. The principal component analysis biplot (Fig. 3d) shows the interdependence between these parameters and examples of multiparameter classification/categorization of IRIFs. The graphs demonstrate that the separation of cell groups of interest—as plotted for tumor vs. tumor-adjacent (normal) tissue cells (left) or cells

allowed to repair DSBs for various PI times (right)—is much better than that solely based on the IRIF numbers. This result demonstrates that the IRIF parameters could be mutually interdependent in a complex way, so their joint consideration may allow categorization of cell groups even in cases when they cannot be separated solely based on the IRIF numbers. In our head and neck squamous cell cancer dataset, the tumor tissue cells and tumor-adjacent tissue cells with similar average IRIF counts per nucleus were distinguished when the morphology of IRIFs (e.g., the average 3D solidity) and intensity of 53BP1 foci were taken into account. DeepFoci analysis also revealed several distinct cell groups within the same dataset, corresponding to cells fixed at different periods of time PI (Fig. 3d, right). In this case, the inclusion of γH2AX focus intensity emerged as an important parameter, in addition to the extent of γH2AX and 53BP1 mutual colocalization. This finding is in accordance with the fact that 53BP1 binds to γH2AX early after DSB induction and dissociates when the damage is repaired, which is also accompanied by γH2AX dephosphorylation. The degree of colocalization between γH2AX and repair proteins thus proved useful for the separation of cells in different phases of the repair process and, in some cases, the separation of normal and tumor cells.

3.8. Single-channel IRIF analysis in multichannel images

In addition to providing analysis based on γH2AX and 53BP1 colocalization, the DeepFoci system enables the quantification of fluorescence channels independently. The applicability of this feature was verified on a dataset composed of permanent cell lines –
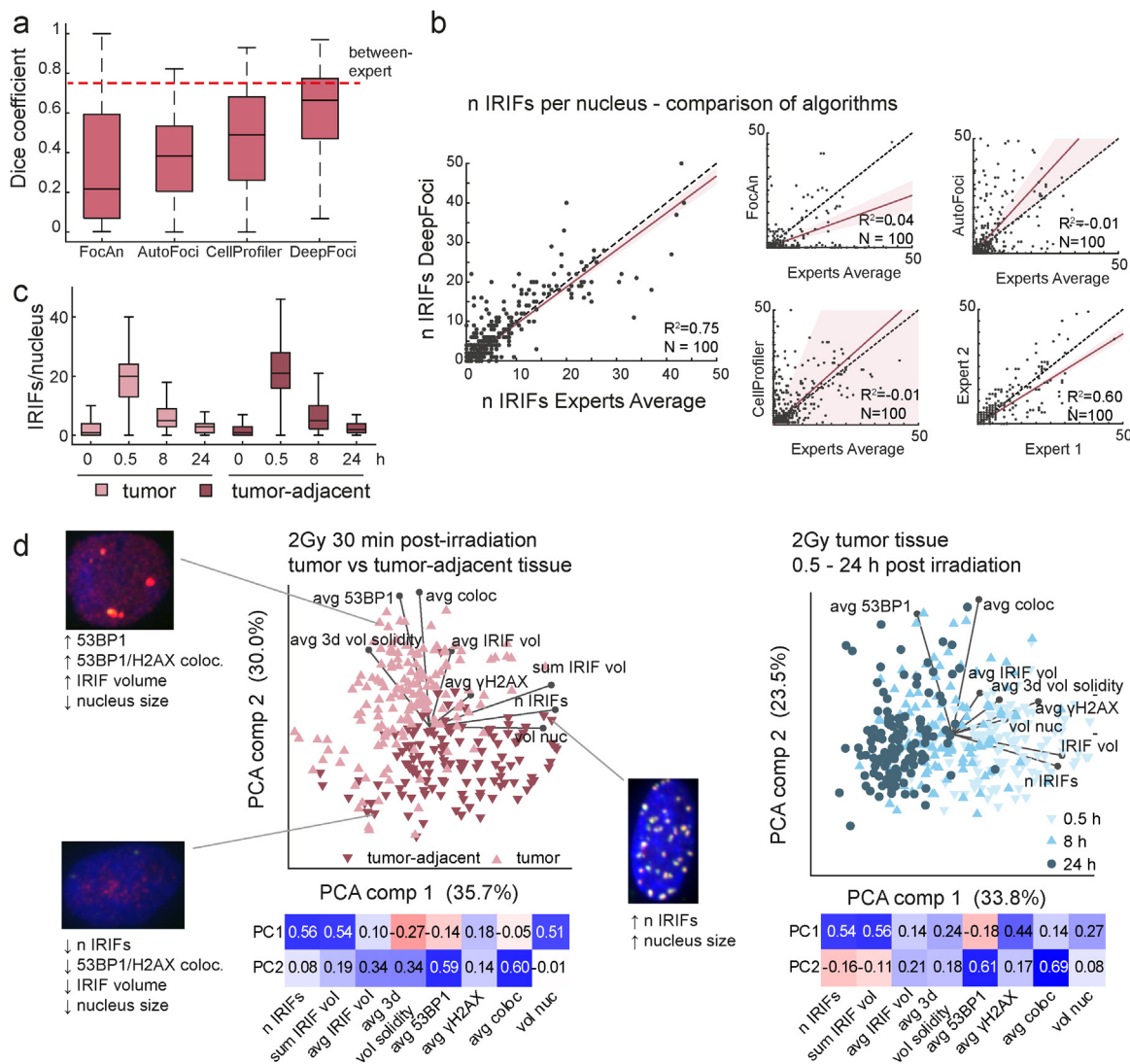
Fig. 3. DeepFoci performance in colocalized IRIF detection. a. Performance comparison for published IRIF-detecting approaches and DeepFoci. The results obtained for the challenging dataset based on head and neck squamous cell cancer primary cultures exposed to 2 Gy 0.5, 8 and 24 h postirradiation are shown. The red dashed line indicates the conformity (Dice coefficient) between experts. **b.** The correlation between automatically and manually detected IRIF numbers compared for DeepFoci and all other tested software methods; the average results for two expert annotations are plotted in the bottom-right graph. **c.** The DSB repair kinetics were determined based on the average IRIF numbers per nucleus detected by DeepFoci at different periods of time (0 min – 24 h) postirradiation. **d.** The principal component analysis biplots show the separation based on γH2AX and 53BP1 IRIF parameters of tumor-adjacent tissue cells and tumor tissues cells (left) and tumor cells fixed at different (0.5, 8 and 24 h) postirradiation times (right). The insets show representative nuclei with IRIF features characteristic of the revealed cell subgroups (categories); the heatmaps below the PCA plots show PCA eigenvectors (negative values in red, positive in blue). Avg - average values for nuclei, n - number of, vol - volume, nuc – nucleus, 53BP1 and γH2AX - average intensities of these stainings, coloc – intensity of colocalization image (multiplication of 53BP1 and gH2AX colocalized IRIFs are used. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

normal human skin fibroblasts (NHDFs) and U-87 glioblastoma cells. U-87 cells are derived from a radioresistant brain tumor that was treated by radiotherapy, while NHDFs are normal (nontransformed) cells with relatively lower radioresistance and are always exposed to radiation during radiotherapy or in the event of a radiation accident. Due to these differences between NHDF and U-87 cells and their different origins, cell-type-specific IRIF morphology and repair dynamics are expected (as already reported in [37]). Moreover, the differences may appear only for one IRIF type (γH2AX or 53BP1, in the present case) or affect both IRIF types in different ways. Indeed, the extent of γH2AX and 53BP1 IRIF colocalization differed between NHDF and U87 cells (see Fig. 5). This difference may be due to different DSB repair capacities and/or repair mechanisms activated in NHDF and U87 cells or as a consequence of repair defects in U87 cancer cells. However, the total number of generated DSBs (labeled by γH2AX) may remain unin-

fluenced. Hence, the estimation of radiation dose based on colocalized γH2AX and 53BP1 (or other repair protein) IRIFs may lead to dose underestimation. Therefore, the regimen of analyzing different IRIFs independently was implemented in DeepFoci.

The cells were exposed to a wide spectrum of doses (0.5–8 Gy) and analyzed at the PI period characteristic of the maximum IRIF occurrence (0.5 h PI) and later (8 h PI) at a stage when most DSBs have already been repaired but the repair of problematic DSBs is still ongoing (see Fig. 5 a). For this evaluation, IRIFs were manually labeled separately in both γH2AX and 53BP1 channels. The same detection network used for the colocalized IRIFs was trained for the two channels. Hence, three networks were trained – two for the detection of γH2AX and 53BP1 IRIFs in individual channels and one for the detection of colocalized IRIFs only (Fig. 4a-c). Following single-channel detection, colocalized foci could be detected from two individually detected channels. Two different approaches
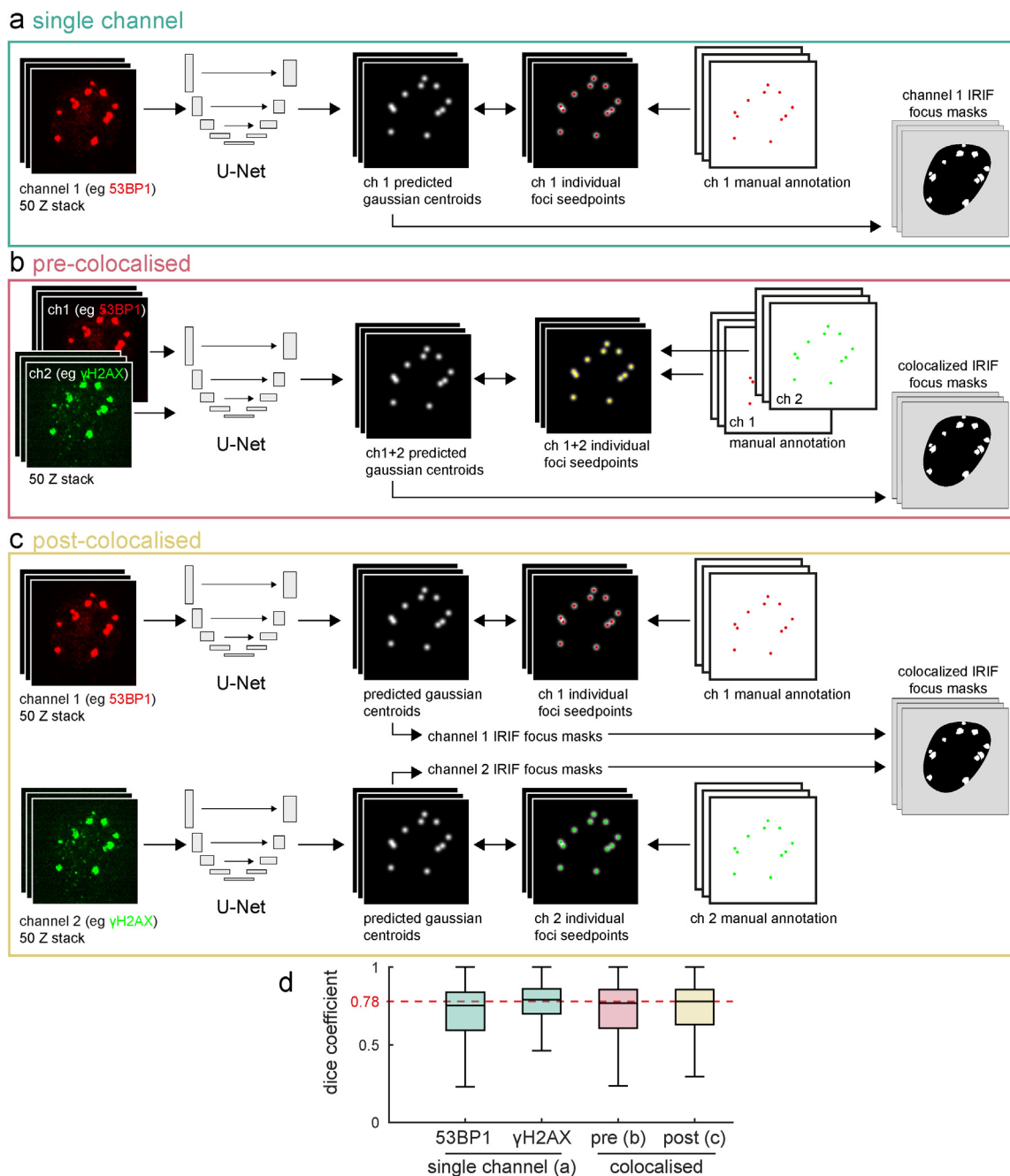
Fig. 4. **Block diagram of single-channel detection for γH2AX and 53BP1 channels in a multichannel image. a.** Single-channel detection of IRIFs using the straightforward application of a detection network trained with single-channel manual labels. **b.** Detection of colocalized IRIFs using labels combined to achieve colocalized labels. The network was trained directly to predict colocalized IRIFs (pre-colocalization). **c.** Detection of colocalized IRIFs using separated training for individual channels and a combination of results (post-colocalization). **d.** Boxplot of Dice coefficients achieved using single-channel detection and detection based on pre/post-colocalized IRIFs (b. and c.) on a dataset with individual IRIFs separately annotated for U87/NHDF cells treated with 0.5–8 Gy 30 min/8h PI (679 FOVs).

were tested: 1) the prediction of colocalized IRIFs trained on merged γH2AX and labeled 53BP1 (pre-colocalization) and 2) the prediction of IRIFs in individual channels by merging the results afterward (post-colocalization) (see Fig. 4 b-c). The colocalization for merging labels/results was defined by a minimal distance of 53BP1 and γH2AX IRIFs and experimentally set to 3.3 μm (20 pixels in x-y). The comparison of the Dice coefficient for single-channel detection and pre/post-colocalization is shown in Fig. 4 d, where post-colocalization achieved slightly better results than pre-colocalization, with median Dice values of 0.780 and 0.769,

respectively. Thus, the post-colocalization approach was used for the following evaluations.

The results were evaluated using 5-fold cross-validation, where the whole experiment (measurement of a specific cell line at a specific time with a specific irradiation dose) was included or excluded for training or testing. Consistent with the manual annotation results, DeepFoci demonstrated stable performance across the different radiation doses, PI times, and cell types, with an overall Dice coefficient of 0.78 (Fig. 5 b-d). A slightly lower accuracy was registered only in nuclei with low IRIF numbers (<10 IRIFs
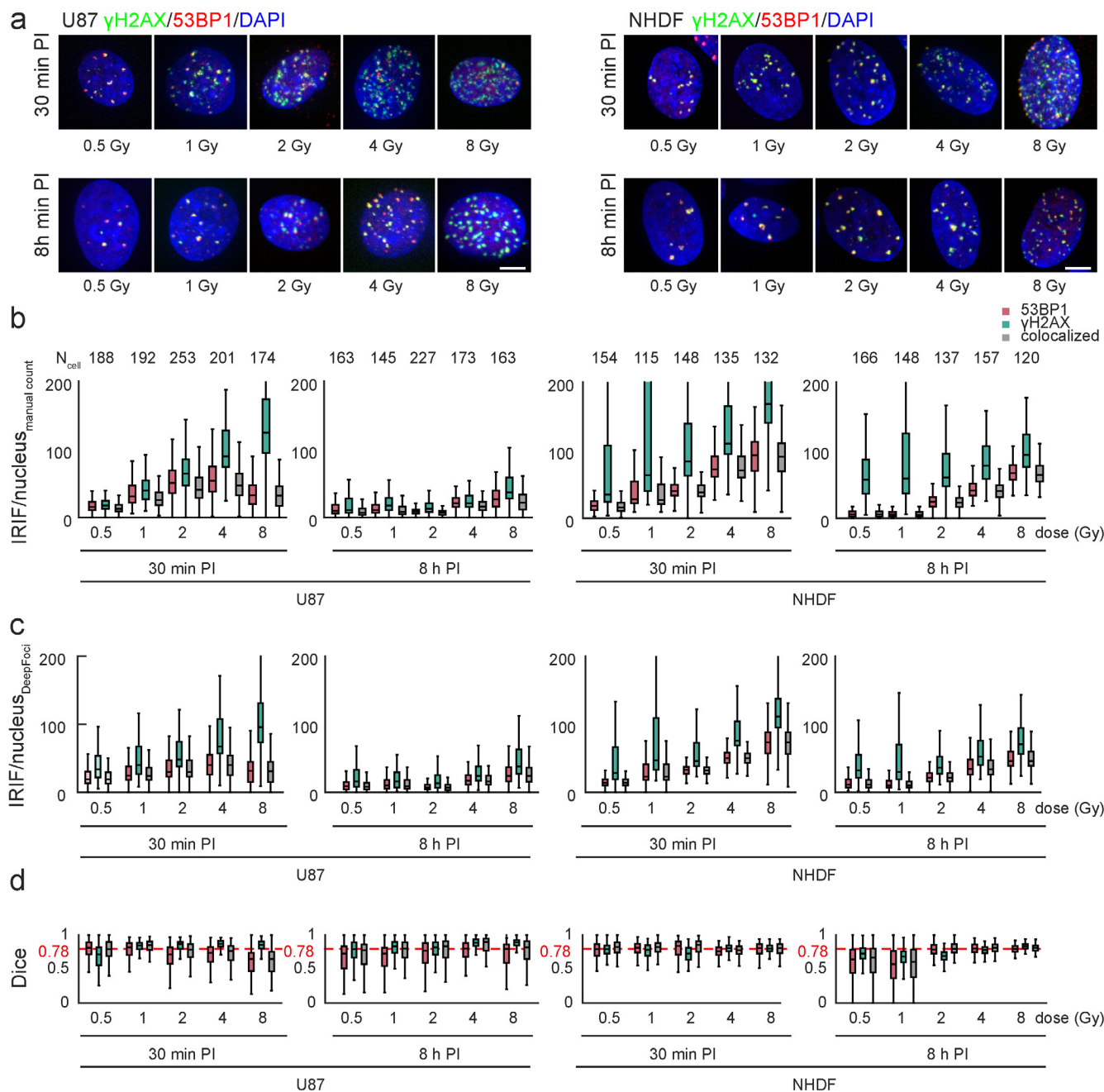
Fig. 5. Automated single-channel detection for γH2AX and 53BP1 in multichannel images compared with manual annotation. **a.** Representative nuclei for different cell lines (normal human skin fibroblasts (NHDFs) and U87 glioblastoma cells), γ-ray doses and times postirradiation. U87 cells represent a radioresistant glioblastoma model, while NHDFs are relatively more radiosensitive. **b.** The distributions of the numbers of γH2AX, 53BP1, and colocalized γH2AX + 53BP1 IRIFs per nucleus obtained by manual annotation (ground truth). **c.** The same distributions as in *b* but obtained by DeepFoci automated analysis. Colocalized IRIFs detected using the approach shown in Fig. 4c. **d.** Software detection accuracy (Dice coefficient). The red dashed line indicates the average Dice coefficient of colocalized IRIFs over all experiments. The boxes show the interquartile range (IQR) with the median indicated; whiskers correspond to maximum and minimum values within 1.5 × IQR. In total, 1412 NHDF and 1879 U87 cells were analyzed. The numbers of cells for individual treatments are shown in the plots of manual annotation (b). Legend: PI, postirradiation; IRIF, ionizing radiation-induced (repair) focus marking DNA double-strand breaks.

per nuclei), specifically NHDF cells exposed to doses lower than 1 Gy (Fig. 5 d).

4. Discussion

The discovery of IRIFs has been a milestone in radiobiology and medical research. Currently, IRIF monitoring represents the most sensitive and versatile method to quantify DSBs and study repair protein interactions and epigenetic modifications at individual DSB sites within the natural environment of the cell nucleus and with time. Immunofluorescence microscopy provides the most complex information on IRIFs, so it has found useful applications in biodosimetry and research. However, robust, precise and reproducible identification of IRIFs still represents an unsatisfactorily solved task, even in the current era of advanced image analysis technologies. *The Second γH2AX-Assay Inter-Comparison Exercise* was carried out in the framework of the European Biodosimetry

Network (RENEB) [30], and our experience has shown that visual (manual) identification of IRIFs still far overperforms any software package in terms of precision.

On the other hand, the manual quantification of IRIFs by a single evaluator is usually unrealistic due to extreme time demands. The cooperation of more evaluators does not help unless they are intensively (i.e., for a long time) trained to reduce interexpert biases in IRIF identification (see Fig. 2f). Equally frustrating is the principal inability of manual analysis to extract IRIF parameters that are of fundamental interest in DNA damage and repair research and that may also be of practical relevance. This precarious situation strongly limits analyses of larger datasets, comparisons of results both within and between laboratories and future progress.

Automated (software) IRIF segmentation is mostly challenged by the tremendous variability of these structures in terms of all their parameters. In particular, the amount of fluorescence can vary greatly between samples. Simple thresholding-based strategies thus provide acceptably accurate and reproducible results only in specific situations, e.g., when the same cell type (e.g., normal lymphocytes) is repeatedly analyzed using a well-optimized staining procedure. Most available methods, therefore, use adaptive thresholding or image standardization; however, this causes the detection sensitivity to be dependent on the number of IRIFs in the nucleus (Fig. 2b). In principle, it remains impossible to set up a threshold parameter universally so that all IRIFs (early, mature, and late) can be correctly recognized and segmented. Specific settings to detect IRIF parameters are often necessary for individual datasets.

According to our experience [13], the difficulty with simple thresholding methods is especially experienced in patient-derived primary cultures, which are characterized by their high heterogeneity. Cells obtained from different patients show, by nature, dramatic differences in IRIF parameters and may even react unpredictably to the same staining protocol (staining procedure optimization for particular patient samples is usually not possible due to a material lack and/or time demands). This was the reason why we included tumor cell primary cultures from different patients in our training and testing datasets.

Particularly problematic are also samples with few or no IRIFs (nonirradiated controls, cells that have already accomplished repair, etc.), in which large numbers of false-positive IRIFs are usually detected as a result of automatic thresholding or image standardization. Of the reports on IRIF detection algorithms, those on Foco [74] or Focinator [71,76] (introduced below) do not disclose the results for control samples. In the paper on FocAn, the controls (time point 0) are included but show unrealistically high IRIF numbers [76].

Several strategies for segmentation of IRIFs (mostly γH2AX or 53BP1 foci) have recently been published [71–75]. Focinator [71,76] is a simple ImageJ macro enabling thresholding, maxima detection, and filtering based on size and circularity. FindFoci [72] is an ImageJ plugin that detects IRIFs as the local maxima. Focus regions are segmented with the downhill gradient algorithm, and the proposed foci are eventually filtered out with specified parameters, which can be trained on a few labeled images. However, the procedure is suitable only for single-channel (γH2AX) labeling so that spatiotemporal interactions between repair proteins or repair proteins and chromatin within the IRIF or cell nucleus cannot be studied. Feng et al. [73] use rather simplistic fuzzy c-means clustering for IRIF detection, which produces noisy and mutually incomparable results if the quantity of IRIFs in nuclei varies to a higher degree. This shortcoming thus seriously complicates even basic analyses of DSB repair kinetics (IRIF number changes with postdamage time), as IRIF numbers may be very high after DNA damage induction (e.g., irradiation) but decrease to zero with repair time. Foco [74] presents an interesting pipeline for nucleus and IRIF segmentation; however, nucleus segmentation is based on intensity thresholding, which we demonstrated to be insufficiently robust for our datasets. AutoFoci [75]—an advanced high-throughput algorithm—extracts several features from each IRIF and finds the most reliable feature to distinguish between the IRIFs and noise. FocAn [76] is the only available 3D IRIF detection method implemented as an ImageJ macro; however, it is based on simple adaptive thresholding followed by maxima detection. CellProfiler offers a universal particle detection algorithm, where customized IRIF detection pipelines can be developed (pipeline from [72] was tested in this paper).

DeepFoci was able to identify and quantify the number of IRIFs with higher accuracy than that of CellProfiler [91], AutoFoci [75] or FocAn [76] and showed precision comparable to that of careful manual analysis performed by a single experienced expert (Fig. 2f, 3a). Accordingly, a consistent and lower level of error was also detected using DeepFoci in cells with lower IRIF counts, which posed serious difficulties to other software packages (Supplementary Fig. 2). The problems with IRIF detection outlined in the previous paragraphs were circumvented by the introduction of a robust 3D segmentation technique based on the standard U-Net. The main modifications involve the application of erosion on segmentation masks and subsequent postprocessing of binary predictions, which lead to the correct separation of individual nuclei. Similarly, the prediction of individual IRIFs via 3D Gaussian functions with specific postprocessing provided precise IRIF detection similar to that with human detection. Additionally, MSER was shown to be a fast and powerful method for the 3D segmentation of IRIFs, producing very precise segmentation results with robust tolerance to different IRIF intensities. With these improvements, the method was proved to achieve satisfactory segmentation of both nuclei and IRIFs. The main advantage of our method is that it is robust against changes in the image intensity. These intensity differences precluded batch analyses of our data, even with the expensive commercial software Imaris. DeepFoci also uses the same U-Net architecture for both nucleus segmentation and IRIF detection, which reduces its implementation complexity. In contrast to manual IRIF counting, the developed method is fast and automatic, and it provides the possibility to extract many other IRIF features in addition to the IRIF count, e.g., mean intensity, size, and solidity. Compared to available automation approaches, DeepFoci is trainable and utilizes advanced deep-learning algorithms. This fundamental advantage leads to better results than those obtained by methods based on thresholding and maxima detection approaches. The proposed method trained on this dataset can also be safely used for a wide range of doses – 0.5 to 8 Gy – without risk of detection accuracy saturation. Moreover, the proposed approach operates on 3D samples. The only other available 3D method is FocAn [76]; however, it utilizes simple IRIF detection approaches and provides unsatisfactory results on our challenging datasets.

The main motivation for this work was to explore whether the obstacles associated with IRIF detection and segmentation in confocal datasets could be overcome by employing deep learning. We aimed to enable an unbiased analysis of large datasets in a timely manner, thereby allowing the realization of complex research studies, effective medical triage (biodosimetry) in the events of mass-casualty radiological incidents, and results comparison between samples and laboratories. In the present manuscript, we introduced DeepFoci, a novel robust software based on deep learning for fully automated identification and morphometric characterization of IRIFs formed by different repair proteins in 3D. The software was designed to overcome current limitations of fluorescence image analysis and allow segmentation of cell nuclei and IRIFs with high fidelity, even in the case of challenging cell specimens of dramatically different quality as they appear in daily

practice. The results confirmed our idea that the precision, specificity and reproducibility of the procedures were significantly enhanced by dual DSB labeling and colocalization analysis of two selected independent DSB markers [21,35]. At the same time, this strategy allowed us to analyze the recruitment of repair proteins into IRIFs and follow their spatiotemporal interactions during repair. These achievements are crucial for both practical (e.g., clinical) and research applications.

It has been well documented (see also Fig. 3e) that individual IRIFs differ quite dramatically in their shapes, sizes, border sharpness and intensities. This variability appears among a) cell types [37], b) cell cultures (especially tumor cell primary cultures) and c) individual cells. In addition, it is problematic to determine and define simple parameters that optimally separate individual IRIFs within IRIF clusters. Using DeepFoci, multiple IRIF parameters were computed, which allowed multiparametric IRIF categorization and thus more accurate recognition of different cell/patient groups and/or repair stages (Fig. 3d). The results indicated that morphometric parameters of IRIFs (such as the 3D solidity) as well as the extent and characteristics of γH2AX and 53BP1 colocalization indeed change between cell types and PI periods. In turn, we show that this information, extracted by DeepFoci, can be used to further refine the identification and categorization of different cell classes or (pre)malignant subclones.

The proposed method was also extended and tested for single-channel detection of IRIFs (detection of IRIFs separately in 53BP1 and γH2AX channels), where it performed consistently across a wide range of radiation doses and different PI times. Detection of IRIFs in individual channels allows the extraction of additional information for different IRIF types in individual nuclei. In a multi-channel setup, this approach also allows the detection of not only colocalized (e.g., 53BP1/γH2AX) IRIFs but also IRIFs that (yet, still or in principle) do not colocalize with IRIFs in the second analyzed channel. Therefore, this regimen also enables analysis of the dynamics of IRIF complex formation and disassembly, which is critical to allow research on repair mechanisms activated by cells at individual DSB sites (see [65] for a review). Moreover, some repair proteins, such as RAD51 (the results for RAD51 are discussed later), that are specific for HR appear almost exclusively in G2 cells and colocalize with only a proportion of γH2AX IRIFs. Hence, dose estimation based on colocalized (e.g.,) γH2AX and RAD51 IRIFs is impossible. Even when a combination of IRIFs that mutually colocalize without any restrictions is used, a lack of colocalization may occur, as observed in the present manuscript for 53BP1 IRIFs, due to saturation of cell repair capacity. Hence, especially for higher radiation doses and short PI times, the detection approach based on IRIF colocalization may lead to severe IRIF and radiation dose underestimation. For the γH2AX and 53BP1 combination, we demonstrate that the degree of colocalization is highly dependent on the dose, PI time, repair protein and cell type. Specifically, the quantity of γH2AX IRIFs was significantly higher than that of 53BP1 IRIFs in all samples but was most prominent in U87 cells. In relatively radiosensitive normal skin fibroblasts (NHDFs), γH2AX and 53BP1 IRIFs were colocalized within most IRIFs; however, distinct saturation was observed for 53BP1 staining in radioresistant U87 glioblastoma cells, especially for doses of 4 Gy and at short (0.5 h) PI periods (Fig. 5b). While this observation may seem paradoxical, delayed and inefficient formation of 53BP1 IRIFs and their lower colocalization with γH2AX in U87 cancer cells than in NHDFs has recently also been confirmed at the nanoscale [37,65], using SMLM [36]. Explanation of this phenomenon is beyond the focus of this manuscript. In general, lower colocalization of γH2AX with 53BP1 at short PI times after high-dose exposures is attributed to a temporary exhaustion of DSB (cNHEJ) repair capacity. The same situation at later PI times may then occur due to ongoing detachment of 53BP1 from DSBs that

were successfully rejoined but where the original chromatin architecture has not yet been restored. The sites of these epimutations [4,92] may then be solely decorated by γH2AX. Noncolocalizing 53BP1 foci, on the other hand, can protect chromosomal fragile sites and DSBs that are transferred unrepaired to the next cell cycle [93]. Hence, DeepFoci broadens the applicability of IRIFs as clinical biomarkers and makes it possible to study DSB repair mechanisms (and their defects) at individual DSB sites.

In addition to observed differences in DSB repair in permanent cell lines, we recognized differences in multiparameter IRIF micromorphology among the tumor cell primary cultures obtained from individual patients suffering from head and neck cancers as well as among the tumor cells and tumor-associated fibroblasts of a particular patient. As head and neck tumors dramatically differ in their radioresistance (even if only HPV(-) tumors are considered), the differences observed herein might be related to this characteristic [13], similar to NHDF and U87 cells. Taken together, in the present manuscript, we show that cell-type differences in γH2AX colocalization with repair proteins and multiparametric differences in the IRIF architecture may indicate alterations in DSB repair in (cancer) cells, which can be observed and evaluated at the microscale using suitable (DeepFoci) software.

Next, the generalizability of the network for other IRIF types and other possible conditions (time PI, cell type, etc.) changes was tested. For this purpose, RAD51 detection was performed with the network trained only for 53BP1 staining (see Supplementary Fig. 3a). RAD51 foci almost exclusively occur in G2 cells and only at DSB sites where the repair operates via HR. Hence, not all γH2AX foci colocalize with RAD51 (even in RAD51-positive cells), and the maximum RAD51 focus occurrence is shifted toward later (4 h PI) periods of time than for 53BP1 (0.5 h PI). The analyses performed at this time point for RAD51 thus tested and confirmed the generalizability of the network with respect to not only the type of IRIF but also the time after irradiation (the network was not trained at this time point, even for 53BP1). The accuracy of detection on a subset of G2 cells (i.e., RAD51-positive cells) was virtually identical to that of the γH2AX and 53BP datasets at different PI times, i.e., the datasets for which the network was specifically trained. The Dice value for RAD51 analysis was 0.76 vs. 0.78 for the analysis of γH2AX and 53BP1 at colocalized channels (see Supplementary Fig. 3b). This result is important even from a practical point of view, as it advocates biodosimetric measurements based on a single DSB marker (e.g., γH2AX) as a less expensive but comparably precise possibility. Hence, the detection based on colocalized channels of two DSB markers could be seemingly redundant or even contradicted by the saturation of some repair proteins described above. However, in real situations, the samples can be collected not earlier than several hours PI, thus eliminating (if it is not too late) the problem of delayed colocalization due to saturation. Moreover, in cells with a high γH2AX or any repair protein background, colocalization-based analysis may still be the best choice.

As also follows from the previous paragraph, it would be of interest in DSB repair studies to correlate the numbers of different IRIF types to the cell cycle phase. This functionality is irrelevant for practical biodosimetry, as lymphocytes, the only cell type relevant for this purpose, are permanently arrested in the G_0 phase of the cell cycle. Moreover, even if nonarrested cell types are used for some specific reason, the easiest approach to analyze how nonsynchronized cells populate in a practical situation is to construct a calibration curve for the whole cell population and exclude only extreme values, e.g., by simply using the median instead of mean values. It should also be noted that in normal cells, in contrast to cancer cells, the background values, even for nonsynchronized cell populations, are extremely low (e.g., [2,14,15,21,79]). Alternatively, in studies where the cells with extreme IRIF values can be expected or are even of special concern, for instance, when the

radiosensitivity/radioresistance of tumors is in question, these extreme cells (likely representing S-phase cells under extensive replication stress [23]) can be separated from the rest of the cell population based on the DNA intensity staining/IRIF numbers, as discussed below. Another approach enabled by our software is to use a DSB marker specific for the S/G2 phase of the cell cycle, with the RAD51 protein used in the present manuscript. This strategy can be used in most types of research studies where specific repair proteins (other than γH2AX and RAD51) are not the subject of research, for instance, when the radiosensitivity and/or repair capacity of cells in different cell cycle phases are compared. Additionally, as the numbers of DSBs generated per Gray depend on the DNA amount, which is doubled in S/G2 cells compared to G1 cells [94–97], the estimations of DSB numbers per radiation dose and DNA unit (Mb) may be specified. However, in some situations, more precise knowledge on the cycle distribution may be critical. When the regulation of DSB repair pathways in the context of the cell cycle is the subject of research, additional markers of the cell cycle are needed, as it cannot be strictly excluded that HR can to some lower extent operate in G1 cells [65]. It would therefore be difficult to judge whether occasional RAD51 foci in some cells represent true foci or staining artifacts.

Given the possibility of estimating the cell cycle by measuring their DNA content according to DNA staining intensity (e.g., as an integrated signal of DAPI in the present study) [94–96], we enriched DeepFoci to be able to extract this parameter. Similar to flow cytometry-based gating, this approach provided DNA content estimations alongside IRIF numbers (see Supplementary Fig. 4). However, it should be noted that the sorting of G1, S and G2 cells based on the integrated DNA signal significantly depends on exactly the same staining and imaging conditions for different samples, and even a weak illumination inhomogeneity within the recorded microscopy field, for instance, might be a problem. Hence, in contrast to flow cytometry, where large numbers of cells are analyzed, giving statistically persuasive data, only some studies have used cell cycle categorization solely based on DNA staining [94–96]. In contrast to these studies but in agreement with other recently reported results [98–101], our datasets did not allow clear separation of the G1 vs. S/G2 cell subpopulation. However, even with the dataset used, the results suggest that DNA staining functionality may serve as a potential filter to remove nuclei with a high DNA content that are therefore atypical (~10% of nuclei in the present study). The uncertainty regarding DNA content as a cell cycle marker for microscopy can be easily overcome by introducing immunofluorescence labeling of a protein with cell cycle-specific expression as an additional color channel. If this strategy is technically impossible, DeepFoci offers, due to its high accuracy of single-channel analysis, the possibility to analyze only one DSB repair protein together with one cell cycle marker. Alternatively, cells specifically modified to express cell phase-specific proteins fused with different GFP spectral variants can be used for experiments (such as Premo FUCCI, Thermo Fisher Scientific, catalog no. P36237), and cells can be synchronized in specific phases of the cell cycle prior to irradiation or sorted by flow cytometry at specific PI periods.

Chromatin architecture and its changes have been shown to play critical roles in DNA damage and during subsequent repair [2,4,50,55,65,79,102–105]. In the present study, the results (not shown) did not confirm systematic changes in chromatin staining intensity at IRIF sites as a function of the radiation dose or PI time. The most likely explanation is that while we studied the PI times starting with 0.5 h PI and ending at 8 h PI, these changes mostly occurred faster after irradiation [79]. Regardless, the ability to quantify DNA staining intensity at sites of IRIFs is a potentially useful functionality of the present software.

Importantly, due to the learning-based nature of the implemented methods, the proposed algorithm offers extensive room for application modifications and can be easily adapted based on the specific requirements of different laboratories, where it can be simply retrained for a different type of data. The evaluation part of the software with the pretrained network contains a very simple and easy-to-use GUI, which is friendly even for an unexperienced user. However, the retraining procedure for new datasets is a complex task and is difficult to cover with a GUI. Therefore, for the retraining, one must run code and, in some cases, make changes in the data-loading part. After retraining both CNNs and readjusting a few parameters (including the focus size range, thresholding step of MSER, and h-minima transform parameter for optimal nucleus separation and focus detection), IRIFs formed by different repair proteins can be analyzed together with their parameters and extent of mutual colocalization in various cell types stained with different methods. Furthermore, the software can be easily tuned to generate comparable results between laboratories for a particular application. This flexibility and robustness, which is highly important, especially for research purposes, represents a unique feature of the introduced software.

5. Conclusion

Quantification of DNA DSBs by means of DSB repair focus (IRIF) immunodetection is of utmost importance in various fields of science and practical life (e.g., medicine, cell biology, radiation protection, space exploration). Because of the nature of IRIFs and fluorescence imaging, where both the IRIF parameters and intensity of analyzed objects may vary dramatically, automatic segmentation of IRIFs and cell nuclei is highly problematic. Accordingly, we developed a new method based on deep learning that overcomes many of the current limitations of image analysis and allows rapid and automated quantification and parameter evaluation of IRIFs. This performance is enabled by a robust U-Net-based technique for nucleus segmentation coupled with U-Net-based focus detection followed by MSER segmentation.

Compared to published approaches, the proposed algorithm works with 3D confocal multichannel data instead of single-channel 2D slices or maximum image projections. This feature makes it possible to extract important additional information on morphological and topological IRIF parameters and not only the focus counts. We believe that the proposed software, whose code is freely available, can substantially simplify DSB quantification and IRIF analysis. Due to the possibility of extracting additional morphometric IRIF and cell nucleus parameters, the software offers numerous practical and research applications. Altogether, the presented software opens the door to a better understanding of IRIF biology and (radiation-induced) DNA damage and repair.

6. Data availability

Data used in the manuscript are publicly available in the Zenodo repository (www.zenodo.com) and on GitHub (www.github.com). Dataset of confocal microscopy of γ-H2AX and 53BP1 DNA repair foci of cells exposed to γ-irradiation – annotated nuclei and IRIFs in primary cultures, pt. 1 (DOI https://doi.org/10.5281/zenodo.4067741) and dose-dependent annotated IRIFs, pt. 2 (DOI https://doi.org/10.5281/zenodo.5549971. MATLAB/Python code for automatic segmentation and for labeling, https://github.com/tomasvicar/DeepFoci).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.11.019.

References

[1] Dickey JS, Redon CE, Nakamura AJ, Baird BJ, Sedelnikova OA, Bonner WM. H2AX: functional roles and potential applications. Chromosoma 2009;118 (6):683–92. https://doi.org/10.1007/s00412-009-0234-4.

[2] Falk M, Lukasova E, Kozubek S. Higher-order chromatin structure in DSB induction, repair and misrepair. Mutat Res 2010;704(1-3):88–100. https://doi.org/10.1016/j.mrrev.2010.01.013.

[3] Falk M, Hausmann M, Lukasova E, Biswas A, Hildenbrand G, Davidkova M, et al. Determining Omics spatiotemporal dimensions using exciting new nanoscopy techniques to assess complex cell responses to DNA damage: part A–radiomics. Crit Rev Eukaryot Gene Expr 2014;24(3):205–23.

[4] Falk M, Hausmann M, Lukášová E, Biswas A, Hildenbrand G, Davídková M, et al. Determining Omics spatiotemporal dimensions using exciting new nanoscopy techniques to assess complex cell responses to DNA damage: part B–structuromics. Crit Rev Eukaryot Gene Expr 2014;24:225–47.

[5] Alhmoud JF, Woolley JF, Al Moustafa A-E, Malki MI. DNA Damage/Repair Management in Cancers. Cancers 2020;12:1050. 10.3390/cancers12041050

[6] Rittich B, Spanová A, Falk M, Benes MJ, Hrubý M. Cleavage of double stranded plasmid DNA by lanthanide complexes. J Chromatogr B Analyt Technol Biomed Life Sci 2004;800:169–73.

[7] Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis: DNA Damage and Repair. Environ Mol Mutagen 2017;58:235–63. https://doi.org/10.1002/em.22087.

[8] Bennett CB, Lewis AL, Baldwin KK, Resnick MA. Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proc Natl Acad Sci 1993;90(12):5613–7. https://doi.org/10.1073/pnas.90.12.5613.

[9] White RR, Vijg J. Do DNA Double-Strand Breaks Drive Aging? Mol Cell 2016;63(5):729–38. https://doi.org/10.1016/j.molcel.2016.08.004.

[10] Madabhushi R, Pan L, Tsai L-H. DNA Damage and Its Links to Neurodegeneration. Neuron 2014;83(2):266–82. https://doi.org/10.1016/j.neuron.2014.06.034.

[11] Gunes S, Al-Sadaan M, Agarwal A. Spermatogenesis, DNA damage and DNA repair mechanisms in male infertility. Reprod Biomed Online 2015;31 (3):309–19. https://doi.org/10.1016/j.rbmo.2015.06.010.

[12] Barnes JL, Zubair M, John K, Poirier MC, Martin FL. Carcinogens and DNA damage. Biochem Soc Trans 2018;46:1213–24. https://doi.org/10.1042/BST20180519.

[13] Falk M, Horakova Z, Svobodova M, Masarik M, Kopecna O, Gumulec J, et al. γH2AX/53BP1 foci as a potential pre-treatment marker of HNSCC tumors radiosensitivity – preliminary methodological study and discussion. Eur Phys J D 2017:241. 10.1140/epjd/e2017-80073-2.

[14] Sevcik J, Falk M, Kleiblova P, Lhota F, Stefancikova L, Janatova M, et al. The BRCA1 alternative splicing variant δ14-15 with an in-frame deletion of part of the regulatory serine-containing domain (SCD) impairs the DNA repair

[15] Sevcik J, Falk M, Macurek L, Kleiblova P, Lhota F, Hojny J, et al. Expression of human BRCA1Δ17-19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response. Cell Signal 2013;25(5):1186–93. https://doi.org/10.1016/j.cellsig.2013.02.008.

[16] Michaelidesová A, Vachelová J, Klementová J, Urban T, Pachnerová Brabcová K, Kaczor S, et al. In vitro comparison of passive and active clinical proton beams. Int J Mol Sci 2020;21(16):5650. https://doi.org/10.3390/ijms21165650.

[17] Burger N, Biswas A, Barzan D, Kirchner A, Hosser H, Hausmann M, et al. A method for the efficient cellular uptake and retention of small modified gold nanoparticles for the radiosensitization of cells. Nanomed Nanotechnol Biol Med 2014;10(6):1365–73. https://doi.org/10.1016/j.nano.2014.03.011.

[18] Štefančíková L, Lacombe S, Salado D, Porcel E, Pagáčová E, Tillement O, et al. Effect of gadolinium-based nanoparticles on nuclear DNA damage and repair in glioblastoma tumor cells. J Nanobiotechnol 2016;14(1). https://doi.org/10.1186/s12951-016-0215-8.

[19] Amarh V, Arthur PK. DNA double-strand break formation and repair as targets for novel antibiotic combination chemotherapy. Future Sci OA 2019;5(8): FSO411. https://doi.org/10.2144/fsoa-2019-0034.

[20] Pagáčová E, Štefančíková L, Schmidt-Kaler F, Hildenbrand G, Vičar T, Depeš D, et al. Challenges and Contradictions of Metal Nano-Particle Applications for Radio-Sensitivity Enhancement in Cancer Therapy. Int J Mol Sci 2019;20 (3):588. https://doi.org/10.3390/ijms20030588.

[21] Hofer M, Falk M, Komůrková D, Falkova I, Bačíková A, Klejdus B, et al. Two New Faces of Amifostine: Protector from DNA Damage in Normal Cells and Inhibitor of DNA Repair in Cancer Cells. J Med Chem 2016;59(7):3003–17. https://doi.org/10.1021/acs.jmedchem.5b0162810.1021/acs.jmedchem.5b01628.s001.

[22] Kratochvílová I, Kopečná O, Bačíková A, Pagáčová E, Falková I, Follett SE, et al. Changes in Cryopreserved Cell Nuclei Serve as Indicators of Processes during Freezing and Thawing. Langmuir ACS J Surf Colloids 2019;35(23):7496–508. https://doi.org/10.1021/acs.langmuir.8b02742.

[23] Falk M, Falková I, Kopečná O, Bačíková A, Pagáčová E, Šimek D, et al. Chromatin architecture changes and DNA replication fork collapse are critical features in cryopreserved cells that are differentially controlled by cryoprotectants. Sci Rep 2018;8(1). https://doi.org/10.1038/s41598-018-32939-5.

[24] Kratochvílová I, Golan M, Pomeisl K, Richter J, Sedláková S, Šebera J, et al. Theoretical and experimental study of the antifreeze protein AFP752, trehalose and dimethyl sulfoxide cryoprotection mechanism: correlation with cryopreserved cell viability. RSC Adv 2017;7(1):352–60. https://doi.org/10.1039/C6RA25095E.

[25] Zahnreich S, Ebersberger A, Kaina B, Schmidberger H. Biodosimetry Based on γ-H2AX Quantification and Cytogenetics after Partial- and Total-Body Irradiation during Fractionated Radiotherapy. Radiat Res 2015;183(4):432. https://doi.org/10.1667/RR13911.110.1667/RR13911.1.s1.

[26] Moquet J, Barnard S, Rothkamm K. Gamma-H2AX biodosimetry for use in large scale radiation incidents: comparison of a rapid '96 well lyse/fix' protocol with a routine method. PeerJ 2014;2:. https://doi.org/10.7717/peerj.282e282.

[27] Jakl L, Marková E, Koláriková L, Belyaev I. Biodosimetry of Low Dose Ionizing Radiation Using DNA Repair Foci in Human Lymphocytes. Genes 2020;11:58. https://doi.org/10.3390/genes11010058.

[28] Viau M, Testard I, Shim G, Morat L, Normil MD, Hempel WM, et al. Global quantification of γH2AX as a triage tool for the rapid estimation of received dose in the event of accidental radiation exposure. Mutat Res Toxicol Environ Mutagen 2015;793:123–31. https://doi.org/10.1016/j.mrgentox.2015.05.009.

[29] Wilkins RC, Carr Z, Lloyd DC. An update of the WHO Biodosenet: Developments since its Inception. Radiat Prot Dosimetry 2016;172(1-3):47–57. https://doi.org/10.1093/rpd/ncw154.

[30] Moquet J, Barnard S, Staynova A, Lindholm C, Monteiro Gil O, Martins V, et al. The second gamma-H2AX assay inter-comparison exercise carried out in the framework of the European biodosimetry network (RENEB). Int J Radiat Biol 2017;93(1):58–64. https://doi.org/10.1080/09553002.2016.1207822.

[31] Staaf E, Brehwens K, Haghdoost S, Czub J, Wojcik A. Gamma-H2AX foci in cells exposed to a mixed beam of X-rays and alpha particles. Genome Integr 2012;3(1):8. https://doi.org/10.1186/2041-9414-3-8.

[32] Cucinotta FA, Durante M. Cancer risk from exposure to galactic cosmic rays: implications for space exploration by human beings. Lancet Oncol 2006;7 (5):431–5. https://doi.org/10.1016/S1470-2045(06)70695-7.

[33] Furukawa S, Nagamatsu A, Nenoi M, Fujimori A, Kakinuma S, Katsube T, et al. Space Radiation Biology for "Living in Space". BioMed Res Int 2020;2020:1–25. https://doi.org/10.1155/2020/4703286.

[34] Jezkova L, Zadneprianetc M, Kulikova E, Smirnova E, Bulanova T, Depes D, et al. Particles with similar LET values generate DNA breaks of different complexity and reparability: a high-resolution microscopy analysis of γH2AX/53BP1 foci. Nanoscale 2018;10(3):1162–79. https://doi.org/10.1039/C7NR06829H.

[35] Depes D, Lee J-H, Bobkova E, Jezkova L, Falkova I, Bestvater F, et al. Single-molecule localization microscopy as a promising tool for γH2AX/53BP1 foci exploration. Eur Phys J D 2018;72. 10.1140/epjd/e2018-90148-1.

[36] Bobkova E, Depes D, Lee J-H, Jezkova L, Falkova I, Pagacova E, et al. Recruitment of 53BP1 Proteins for DNA Repair and Persistence of Repair

Clusters Differ for Cell Types as Detected by Single Molecule Localization Microscopy. Int J Mol Sci 2018;19(12):3713. https://doi.org/10.3390/ijms19123713.

[37] Hausmann M, Neitzel C, Bobkova E, Nagel D, Hofmann A, Chramko T, et al. Single Molecule Localization Microscopy Analyses of DNA-Repair Foci and Clusters Detected along Particle Damage Tracks. Front Phys Sect Med Phys. Imaging 2020;8. https://doi.org/10.3389/fphy.2020.578662.

[38] Rogakou EP, Boon C, Redon C, Bonner WM. Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol 1999;146:905–16. https://doi.org/10.1083/jcb.146.5.905.

[39] Firsanov DV, Solovjeva LV, Svetlova MP. H2AX phosphorylation at the sites of DNA double-strand breaks in cultivated mammalian cells and tissues. Clin Epigenetics 2011;2(2):283–97. https://doi.org/10.1007/s13148-011-0044-4.

[40] Redon CE, Dickey JS, Bonner WM, Sedelnikova OA. γ-H2AX as a biomarker of DNA damage induced by ionizing radiation in human peripheral blood lymphocytes and artificial skin. Adv Space Res 2009;43(8):1171–8. https://doi.org/10.1016/j.asr.2008.10.011.

[41] Mariotti LG, Pirovano G, Savage KI, Ghita M, Ottolenghi A, Prise KM, et al. Use of the γ-H2AX assay to investigate DNA repair dynamics following multiple radiation exposures. PLoS ONE 2013;8(11):e79541.

[42] Durdik M, Kosik P, Gursky J, Vokalova L, Markova E, Belyaev I. Imaging flow cytometry as a sensitive tool to detect low-dose-induced DNA damage by analyzing 53BP1 and γH2AX foci in human lymphocytes: Imaging Flow Cytometry for DNA Damage Analysis. Cytometry A 2015;87(12):1070–8. https://doi.org/10.1002/cyto.a.22731.

[43] Sharma A, Singh K, Almasan A. Histone H2AX Phosphorylation: A Marker for DNA Damage. In: Bjergbæk L, editor. DNA Repair Protoc, vol. 920. Totowa, NJ: Humana Press; 2012. p. 613–26. https://doi.org/10.1007/978-1-61779-998-3_40.

[44] Lee Y, Wang Q, Shuryak I, Brenner DJ, Turner HC. Development of a high-throughput γ-H2AX assay based on imaging flow cytometry. Radiat Oncol 2019;14:150. https://doi.org/10.1186/s13014-019-1344-7.

[45] Takahashi A, Ohnishi T. Does γH2AX foci formation depend on the presence of DNA double strand breaks? Cancer Lett 2005;229(2):171–9. https://doi.org/10.1016/j.canlet.2005.07.016.

[46] Ceelen M, van Weissenbruch MM, Vermeiden JPW, van Leeuwen FE, Delemarre-van de Waal HA. Growth and development of children born after in vitro fertilization. Fertil Steril 2008;90(5):1662–73. https://doi.org/10.1016/j.fertnstert.2007.09.005.

[47] Her J, Bunting SF. How cells ensure correct repair of DNA double-strand breaks. J Biol Chem 2018;293(27):10502–11. https://doi.org/10.1074/jbc.TM118.000371.

[48] Scherthan H, Lee J-H, Maus E, Schumann S, Muhtadi R, Chojowski R, et al. Nanostructure of Clustered DNA Damage in Leukocytes after In-Solution Irradiation with the Alpha Emitter Ra-223. Cancers 2019;11. 10.3390/cancers11121877.

[49] Bach M, Savini C, Krufczik M, Cremer C, Rösl F, Hausmann M. Super-Resolution Localization Microscopy of γ-H2AX and Heterochromatin after Folate Deficiency. Int J Mol Sci 2017;18(8):1726. https://doi.org/10.3390/ijms18081726.

[50] Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: a historical perspective. Nat Rev Cancer 2016;16(1):35–42. https://doi.org/10.1038/nrc.2015.4.

[51] Misteli T, Soutoglou E. The emerging role of nuclear architecture in DNA repair and genome maintenance. Nat Rev Mol Cell Biol 2009;10(4):243–54. https://doi.org/10.1038/nrm2651.

[52] Jakob B, Rudolph JH, Gueven N, Lavin MF, Taucher-Scholz G. Live cell imaging of heavy-ion-induced radiation responses by beamline microscopy. Radiat Res 2005;163(6):681–90.

[53] Kruhlak MJ, Celeste A, Dellaire G, Fernandez-Capetillo O, Müller WG, McNally JG, et al. Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. J Cell Biol 2006;172:823–34. 10.1083/jcb.200510015.

[54] Noon AT, Goodarzi AA. 53BP1-mediated DNA double strand break repair: Insert bad pun here. DNA Repair 2011;10(10):1071–6. https://doi.org/10.1016/j.dnarep.2011.07.012.

[55] Goodarzi AA, Jeggo P, Lobrich M. The influence of heterochromatin on DNA double strand break repair: Getting the strong, silent type to relax. DNA Repair 2010;9(12):1273–82. https://doi.org/10.1016/j.dnarep.2010.09.013.

[56] Reindl J, Girst S, Walsh DWM, Greubel C, Schwarz B, Siebenwirth C, et al. Chromatin organization revealed by nanostructure of irradiation induced γH2AX, 53BP1 and Rad51 foci. Sci Rep 2017;7(1). https://doi.org/10.1038/srep40616.

[57] Rothkamm K, Horn S. gamma-H2AX as protein biomarker for radiation exposure. Ann Ist Super Sanita 2009;45:265–71.

[58] Rybak P, Hoang A, Bujnowicz L, Bernas T, Berniak K, Zarębski M, et al. Low level phosphorylation of histone H2AX on serine 139 (γH2AX) is not associated with DNA double-strand breaks. Oncotarget 2016;7(31):49574–87.

[59] Schneider J, Weiss R, Ruhe M, Jung T, Roggenbuck D, Stohwasser R, et al. Open source bioimage informatics tools for the analysis of DNA damage and associated biomarkers. J Lab Precis Med 2019(4):21. https://doi.org/10.21037/jlpm10.21037/jlpm.2019.04.05.

[60] Fu Q, Wang J, Huang T. Characterizing the DNA damage response in fibrosarcoma stem cells by in-situ cell tracking. Int J Radiat Biol 2019;95(2):99–106. https://doi.org/10.1080/09553002.2019.1539879.

[61] Rothkamm K, Barnard S, Moquet J, Ellender M, Rana Z, Burdak-Rothkamm S. DNA damage foci: Meaning and significance. Environ Mol Mutagen 2015;56(6):491–504. https://doi.org/10.1002/em.v56.610.1002/em.21944.

[62] Barnard S, Ainsbury EA, Al-hafidh J, Hadjidekova V, Hristova R, Lindholm C, et al. The first gamma-H2AX biodosimetry intercomparison exercise of the developing European biodosimetry network RENEB. Radiat Prot Dosimetry 2015;164(3):265–70. https://doi.org/10.1093/rpd/ncu259.

[63] Einbeck J, Ainsbury EA, Sales R, Barnard S, Kaestle F, Higueras M, et al. A statistical framework for radiation dose estimation with uncertainty quantification from the γ-H2AX assay. PLoS ONE 2018;13(11):e0207464.

[64] Falk M, Hausmann M. A Paradigm Revolution or Just Better Resolution – Will Newly Emerging Superresolution Techniques Identify Chromatin Architecture as a Key Factor in Radiation-Induced DNA Damage and Repair Regulation? Cancers 2021;13:article 18 (1-30). doi.org/10.3390/cancers13010018.

[65] Hahn H, Neitzel Charlotte, Kopečná Olga, Heermann W Dieter, Falk Martin, Hausmann A. Topological analysis of γH2AXand MRE11 clusters detected by localization microscopy during X-ray induced DNA double- strand break repair. Cancers 2021;13(21): article 5561; 10.3390/cancers13215561.

[66] Falk M, Hausmann M. Advances in research of DNA damage and repair in cells exposed to various types of ionizing radiation in the era of super-resolution optical microscopy. Cas Lek Cesk 2020.

[67] Willers H, Taghian AG, Luo C-M, Treszezamsky A, Sgroi DC, Powell SN. Utility of DNA Repair Protein Foci for the Detection of Putative BRCA1 Pathway Defects in Breast Cancer Biopsies. Mol Cancer Res 2009;7(8):1304–9. https://doi.org/10.1158/1541-7786.MCR-09-0149.

[68] Bonner WM, Redon CE, Dickey JS, Nakamura AJ, Sedelnikova OA, Solier S, et al. GammaH2AX and cancer. Nat Rev Cancer 2008;8:957–67. https://doi.org/10.1038/nrc2523.

[69] Anglada T, Repullés J, Espinal A, LaBarge MA, Stampfer MR, Genescà A, et al. Delayed γH2AX foci disappearance in mammary epithelial cells from aged women reveals an age-associated DNA repair defect. Aging 2019;11(5):1510–23.

[70] Oeck S, Malewicz NM, Hurst S, Rudner J, Jendrossek V. The Focinator - a new open-source tool for high-throughput foci evaluation of DNA damage. Radiat Oncol Lond Engl 2015;10:163. https://doi.org/10.1186/s13014-015-0453-1.

[71] Herbert AD, Carr AM, Hoffmann E, Lichten M. FindFoci: a focus detection algorithm with automated parameter training that closely matches human assignments, reduces human inconsistencies and increases speed of analysis. PLoS ONE 2014;9(12):e114749.

[72] Feng J, Lin J, Zhang P, Yang S, Sa Y, Feng Y. A novel automatic quantification method for high-content screening analysis of DNA double strand-break response. Sci Rep 2017;7:9581. https://doi.org/10.1038/s41598-017-10063-0.

[73] Lapytsko A, Kollarovic G, Ivanova L, Studencka M, Schaber J. FoCo: a simple and robust quantification algorithm of nuclear foci. BMC Bioinf 2015;16:392. https://doi.org/10.1186/s12859-015-0816-5.

[74] Lengert N, Mirsch J, Weimer RN, Schumann E, Haub P, Drossel B, et al. AutoFoci, an automated high-throughput foci detection approach for analyzing low-dose DNA double-strand break repair. Sci Rep 2018;8(1). https://doi.org/10.1038/s41598-018-35660-5.

[75] Oeck S, Malewicz NM, Hurst S, Al-Refae K, Krysztofiak A, Jendrossek V. The Focinator v2–0 – Graphical Interface, Four Channels, Colocalization Analysis and Cell Phase Identification. Radiat Res 2017;188(1):114–20. https://doi.org/10.1667/RR14746.1.

[76] Memmel S, Sisario D, Zimmermann H, Sauer M, Sukhorukov VL, Djuzenova CS, et al. FocAn: automated 3D analysis of DNA repair foci in image stacks acquired by confocal fluorescence microscopy. BMC Bioinf 2020;21(1). https://doi.org/10.1186/s12859-020-3370-8.

[77] Svobodova M, Raudenska M, Gumulec J, Balvan J, Fojtu M, Kratochvilova M, et al. Establishment of oral squamous cell carcinoma cell line and magnetic bead-based isolation and characterization of its CD90/ CD44 subpopulations. Oncotarget 2017;8(39):66254–69.

[78] Falk M, Lukasova E, Gabrielova B, Ondrej V, Kozubek S. Chromatin dynamics during DSB repair. Biochim Biophys Acta 2007;1773(10):1534–45. https://doi.org/10.1016/j.bbamcr.2007.07.002.

[79] Matula P, Maška M, Daněk O, Matula P, Kozubek M. Acquiarium: Free software for the acquisition and analysis of 3D images of cells in fluorescence microscopy. Proc - IEEE Int Symp Biomed Imag Nano Macro ISBI 2009:1138–41. https://doi.org/10.1109/ISBI.2009.5193258.

[80] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Trans Pattern Anal Mach Intell 2012;34(11):2274–82. https://doi.org/10.1109/TPAMI.2012.120.

[81] Ulman V, Maška M, Magnusson KEG, Ronneberger O, Haubold C, Harder N, et al. An objective comparison of cell-tracking algorithms. Nat Methods 2017;14(12):1141–52. https://doi.org/10.1038/nmeth.4473.

[82] Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. Transl Cancer Res 2013;2:130–43. https://doi.org/10.3978/j.issn.2218-676X.2013.04.02.

[83] Ensminger M, Löbrich M. One end to rule them all: Non-homologous end-joining and homologous recombination at DNA double-strand breaks. Br J Radiol 2020;93(1115):20191054. https://doi.org/10.1259/bjr.20191054.

[84] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Med. Image Comput. Comput.-Assist. Interv. – MICCAI

2015, vol. 9351, Cham: Springer International Publishing; 2015, p. 234–41. 10.1007/978-3-319-24574-4_28.

[85] Meyer F, Beucher S. Morphological segmentation. J Vis Commun Image Represent 1990;1(1):21–46. https://doi.org/10.1016/1047-3203(90)90014-M.

[86] Matas J, Chum O, Urban M, Pajdla T. Robust wide-baseline stereo from maximally stable extremal regions. Image Vis Comput 2004;22(10):761–7. https://doi.org/10.1016/j.imavis.2004.02.006.

[87] Parvati K, Prakasa Rao BS, Mariya DM. Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation. Discrete Dyn Nat Soc 2008;2008:1–8. https://doi.org/10.1155/2008/384346.

[88] Vedaldi A, Fulkerson B. Vlfeat: an open and portable library of computer vision algorithms. Proc. Int. Conf. Multimed. - MM 10, Firenze, Italy: ACM Press; 2010, p. 1469. 10.1145/1873951.1874249.

[89] Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In: Cardoso MJ, Arbel T, Carneiro G, Syeda-Mahmood T, Tavares JMRS, Moradi M, et al., editors. Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support, vol. 10553, Cham: Springer International Publishing; 2017, p. 240–8. 10.1007/978-3-319-67558-9_28.

[90] Jones TR, Kang IH, Wheeler DB, Lindquist RA, Papallo A, Sabatini DM, et al. Cell Profiler Analyst: Data exploration and analysis software for complex image-based screens. BMC Bioinf 2008;9(1). https://doi.org/10.1186/1471-2105-9-482.

[91] Ježková L, Falk M, Falková I, Davídková M, Bačíková A, Štefančíková L, et al. Function of chromatin structure and dynamics in DNA damage, repair and misrepair: γ-rays and protons in action. Appl Radiat Isot Data Instrum Methods Use Agric Ind Med 2014;83 Pt B:128–36. 10.1016/j.apradiso.2013.01.022.

[92] Lukas C, Savic V, Bekker-Jensen S, Doil C, Neumann B, Sølvhøj Pedersen R, et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat Cell Biol 2011;13(3):243–53. https://doi.org/10.1038/ncb2201.

[93] Kruszewski M, Iwanenko T, Machaj EK, Oldak T, Wojewodzka M, Kapka-Skrzypczak L, et al. Direct use of the comet assay to study cell cycle distribution and its application to study cell cycle-dependent DNA damage formation. Mutagenesis 2012;27(5):551–8. https://doi.org/10.1093/mutage/ges018.

[94] Roukos V, Pegoraro G, Voss TC, Misteli T. Cell cycle staging of individual cells by fluorescence microscopy. Nat Protoc 2015;10(2):334–48. https://doi.org/10.1038/nprot.2015.016.

[95] Löbrich M, Shibata A, Beucher A, Fisher A, Ensminger M, Goodarzi AA, et al. γH2AX foci analysis for monitoring DNA double-strand break repair: Strengths, limitations and optimization. Cell Cycle 2010;9(4):662–9. https://doi.org/10.4161/cc.9.4.10764.

[96] Beucher A, Birraux J, Tchouandong L, Barton O, Shibata A, Conrad S, et al. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. EMBO J 2009;28(21):3413–27. https://doi.org/10.1038/emboj.2009.276.

[97] Hsu M-Y, Yang MH, Schnegg CI, Hwang S, Ryu B, Alani RM. Notch3 signaling-mediated melanoma–endothelial crosstalk regulates melanoma stem-like cell homeostasis and niche morphogenesis. Lab Invest 2017;97(6):725–36. https://doi.org/10.1038/labinvest.2017.1.

[98] Kashyap M, Das D, Preet R, Mohapatra P, Satapathy SR, Siddharth S, et al. Scaffold hybridization in generation of indenoindolones as anticancer agents that induce apoptosis with cell cycle arrest at G2/M phase. Bioorg Med Chem Lett 2012;22(7):2474–9. https://doi.org/10.1016/j.bmcl.2012.02.007.

[99] Campbell JM, Habibalahi A, Mahbub S, Gosnell M, Anwer AG, Paton S, et al. Non-destructive, label free identification of cell cycle phase in cancer cells by multispectral microscopy of autofluorescence. BMC Cancer 2019;19(1). https://doi.org/10.1186/s12885-019-6463-x.

[100] Lopez Perez R, Münz F, Kroschke J, Brauer J, Nicolay NH, Huber PE. Cell Cycle-specific Measurement of & #947;H2AX and Apoptosis After Genotoxic Stress by Flow Cytometry. J Vis Exp 2019;59968. https://doi.org/10.3791/59968.

[101] Falk M, Lukasova E, Gabrielova B, Ondrej V, Kozubek S. Local changes of higher-order chromatin structure during DSB-repair. J Phys Conf Ser 2008;101:012018. https://doi.org/10.1088/1742-6596/101/1/012018.

[102] Falk M, Lukásová E, Kozubek S. Chromatin structure influences the sensitivity of DNA to gamma-radiation. Biochim Biophys Acta 2008;1783:2398–414. https://doi.org/10.1016/j.bbamcr.2008.07.010.

[103] Goodarzi AA, Noon AT, Deckbar D, Ziv Y, Shiloh Y, Löbrich M, et al. ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. Mol Cell 2008;31(2):167–77. https://doi.org/10.1016/j.molcel.2008.05.017.

[104] Jakob B, Splinter J, Conrad S, Voss K-O, Zink D, Durante M, et al. DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. Nucleic Acids Res 2011;39:6489–99. 10.1093/nar/gkr230.

[105] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Med. Image Comput. Comput.-Assist. Interv. – MICCAI 2016, vol. 9901, Cham: Springer International Publishing; 2016, p. 424–32. 10.1007/978-3-319-46723-8_49.