Proceedings of the 11th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'25)
Paris, France – August, 2025
Paper No. ICBES 133
DOI: 10.11159/icbes25.133

Machine Learning-Driven Image Processing for Investigating DNA Repair after Damage and Calcium Responses in Cellular Injury Models

Connor Lee[1, †], Harvy Chang[1, †], Hongyi Niu[1, †], Albert Li[1, †], Chengbiao Wu[2], Veronica Gomez-Godinez[1], Linda Shi[1+]

[1]Institute of Engineering in Medicine, University of California San Diego, La Jolla, CA, 92093
[2]Department of Neuroscience Biology, University of California San Diego, La Jolla, CA,92093
***** Correspondence: zshi@ucsd.edu
†High school students participating in IEM OPALS program

Abstract **-** This paper presents the development and application of machine learning (ML)-based image processing pipelines to investigate DNA repair and calcium responses following cellular injury. We induced controlled DNA double-strand breaks through laser ablation and also simulated traumatic injuries in live-cell models using a similar method of laser-induced shockwave (LIS) systems. By integrating Cellpose-based segmentation and Python-driven automation, we substantially improved the image data analysis of three major applications: protein recruitment after DNA damage, calcium flux tracking in cortical neurons after shockwave injury, and calcium dynamics comparison in Alzheimer's disease (AD) models. We considerably shortened manual processing time while maintaining improved levels of precision, as well as being easily scalable to other applications. These results demonstrate the potential of ML-enhanced image analysis in advancing research of DNA damage repair, traumatic brain injury (TBI) simulation, and neurodegeneration studies.

Keywords: Machine Learning, Image Processing, DNA Damage Repair, Laser-Induced Shockwave, Calcium Imaging, Alzheimer's Disease, Cellpose, Automated Analysis

1. Introduction

The elucidation of the DNA double helix structure by Watson and Crick significantly advanced our understanding of genetic material, subsequently highlighting DNA's susceptibility to damage. Early seminal work by Meselson and Stahl established foundational principles in DNA replication and repair mechanisms [1]. In the 1970s, Berns and colleagues used precise laser technology to introduce laser ablation as a method to generate cellular-level damage in a highly controlled manner [2], which greatly enhanced the study of biological damage processes [3]. Further advancements in laser technologies and live imaging techniques have enabled real-time monitoring of DNA repair processes within living cells, allowing for a better understanding of molecular interactions following DNA double-strand breaks (DSBs) [4-11].

Despite these technological advancements, analyzing the large amounts of imaging data generated from live-cell experiments remains labor-intensive and highly error-prone. Consequently, there is a significant need for efficient, automated data analysis methods. The integration of machine learning (ML) algorithms offers a powerful solution to this challenge. In this paper, we proposed an ML-driven image analysis approach that substantially accelerates data processing while delivering accuracy and being highly reproducible.

Additionally, laser-induced shockwaves (LIS) have been serving as a powerful tool for studying the cellular damage and repair processes induced by traumatic brain injury (TBI), driven by its prevalence in sports and military contexts. Research in blast-induced TBI began in the 1950s [12], evolving into extensive molecular studies over subsequent decades [13-14]. Inspired by applications of LIS in medical treatments such as lithotripsy of gallstones [15], recent research has explored the biological effects of therapeutic shockwaves and ultrasound in neural tissues, explicitly emphasizing the role of Ca2+ signaling mediated by cavitation microbubbles [16]. Our laboratory has developed a specialized LIS system capable of precisely simulating cellular conditions analogous to blast-induced TBI, allowing for the detailed examination of neuronal and astrocytic calcium responses [17].

Traditional manual quantification methods for calcium fluctuations are inefficient and susceptible to variability. We integrated advanced LIS techniques with robust Python-based machine learning methods for rapid, precise, and automated analysis of neuronal calcium dynamics, facilitating comparisons between healthy and Alzheimer's disease models. Through this method, we aim to enhance the efficiency and reliability of biological image analysis, thereby advancing our understanding of DNA repair mechanisms and calcium signaling following cellular injury.

2. Materials and System Setups

2.1 Laser Ablation System Setup and Cell Lines for DNA Repair Study

A tunable femtosecond mode-locked Ti:Sapphire infrared laser (Mai Tai, Spectra-Physics, Newport Corp., Mountain View, CA) was used to generate the laser-induced microirradiation for precise subcellular targeting. Laser power was attenuated via a motorized rotating optical polarizer (Newport, Irvine, CA), and pulse delivery was controlled by a mechanical shutter (Vincent Associates, Rochester, NY) with a 10 ms duty cycle. The laser beam was expanded to fill the back aperture of a 100× NA 1.3 Zeiss objective mounted on a Zeiss Axiovert 200M microscope and focused on the sample. Two-photon excitation was employed at either 730 nm (effective 365 nm) or 800 nm (effective 400 nm) with laser powers of 50 mW and 60 mW, respectively, measured before entering the phase contrast objective. A custom-built Labview APP was programmed to facilitate high-resolution, real-time, live-cell imaging.

The preparation of U2OS (human osteosarcoma) cell lines was described previously in [18]. U2OS cells were obtained from the American Type Culture Collection (ATCC) cell repository. RPE-1 WT cells were received from Dr. Stephen P. Jackson's lab. UWB1 and UWB1 reconstituted with BRCA1 cells were received from Dr. Lee Zou's lab. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal bovine serum (FBS; GeminiBio.), 2mM L-glutamine (Sigma-Aldrich), and 1% penicillin-streptomycin containing glutamine (Gibco) at 37 °C in a humid atmosphere containing 5% CO_2. The U2OS (EGFP-HR/STGC) reporter cell line was generated by the transfection of the EGFP-HR/STGC reporter into U2OS cells with polyethylenimine (PEI) using the standard protocol, followed by hygromycin B (100 μg/ml) selection. Double-strand breaks (DSBs) were introduced in live-cell nuclei through precision laser-induced microirradiation.

2.2 Laser-Induced Shockwave System Setup

A Coherent Flare 532 nm laser system (100 Hz repetition rate, 2 ns pulse width, 450 μJ pulse energy; Spectra-Physics, Mountain View, CA) was used to generate localized shockwaves. Laser power was attenuated via a motorized rotating optical polarizer (Newport, Irvine, CA), and pulse delivery was controlled by a mechanical shutter (Vincent Associates, Rochester, NY) with a 10–15 ms duty cycle. The laser beam was expanded to fill the back aperture of a 40× NA 1.3 Zeiss objective mounted on a Zeiss 200M microscope and focused 10 μm above the substrate. The incident power measured before the objective was approximately 200–220 μW. A Zeiss filter set 48 (436/20 nm excitation, 455 nm long-pass dichroic mirror) was installed, with additional emission filters (535/30 nm for FRET, 480/40 nm for ECFP) mounted on a LUDL motorized filter wheel, positioned before an ORCA-Flash4.0 V2 Hamamatsu CMOS camera for fluorescence imaging.

Mouse primary cortical neurons were cultured as described in (Gu et al., DOI: 10.1016/j.nbd.2024.106502). Cells were preloaded with Fluo-4 AM[17] to study calcium flux. LIS creates a cavitation bubble that expands and rapidly collapses, causing subsequent death to cells in the bubble's vicinity. A 1032 nm laser was focused 10 μm above the substrate, which contained the neurons.

3. Results and Discussions

3.1 Machine Learning-Driven Image Analysis for DNA Repair Dynamics Post-Laser Ablation

One of the critical bottlenecks traditionally faced in live-cell DNA repair imaging is the labor-intensive nature of manually analyzing hundreds to thousands of time-lapse fluorescence images, capturing subtle recruitment events of repair proteins post-laser ablation. To overcome this challenge and significantly expedite the data analysis phase, we designed and adapted a robust machine learning pipeline, building on the strengths of Cellpose [19], a state-of-the-art deep learning segmentation framework in computational biology.

Our work represents a major step forward by integrating automated cell detection, repair line identification, and intensity quantification into a single, streamlined system. After extensively evaluating many of the available methods through online repositories, published studies, and GitHub resources, we found Cellpose to be the most versatile. We customized Cellpose's pre-trained convolutional neural networks (CNNs) to specifically address the challenges associated with dynamic protein recruitment analysis following DNA damage.

The pipeline we developed is as follows:

1. **Batch Image Loading**: Sequential fluorescence images capturing the recruitment of repair proteins, such as Red52, were loaded (e.g., Figure 1(a)).
2. **Automated Cell Segmentation**: Leveraging deep learning algorithms within Cellpose, each image underwent automated segmentation to detect cellular boundaries and generate accurate masks (e.g., Figure 1(b)).
3. **Cropped Region Extraction**: Using segmentation masks, individual cells were cropped and saved for high-precision analysis.
4. **Recruitment Line Detection Initiation**: The brightest pixel within each cropped cell was identified using np.argmax() in Python, pinpointing the initial recruitment signal at laser cut sites.
5. **Thresholding and Binary Mask Creation**: The cropped images were normalized, thresholded, and converted to binary masks using OpenCV libraries, enhancing the signal-to-noise ratio.
6. **Contour Detection and Line Segmentation**: Contours were detected, and the laser-induced cut line was precisely delineated by finding and connecting the two farthest points on the contour, as shown in Figure 2(a).
7. **Quantitative Brightness Measurement**: A quantitative analysis was performed by calculating the mean fluorescence intensity along the laser cut line, providing a direct measurement of recruitment kinetics.
8. **Data Export and Visualization**: Results, including brightness values over time, were systematically recorded into CSV files and plotted using Matplotlib to generate graphical representations of recruitment dynamics (e.g., Figure 2(b) of the second cell in Figure 2(a)).

This machine learning-driven approach not only speeds up the data analysis by reducing processing times significantly compared to manual methods but also enhances accuracy and eliminates subjective bias that is inherent to traditional image scoring. The successful adaptation of Cellpose and the custom pipeline serve as a highly scalable model for the broader biological community studying DNA repair mechanisms, setting a new benchmark for automated, efficient biological imaging analysis.

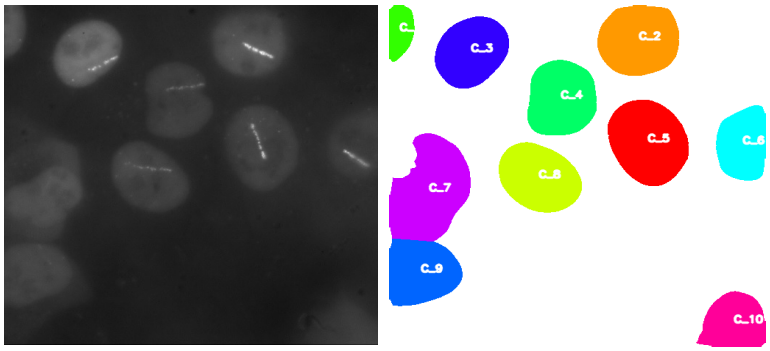


Figure 1(a): Post-2 min Image of U2OS Red52 Protein Recruitment, **(b):** Automated Cell Detection Using Cellpose

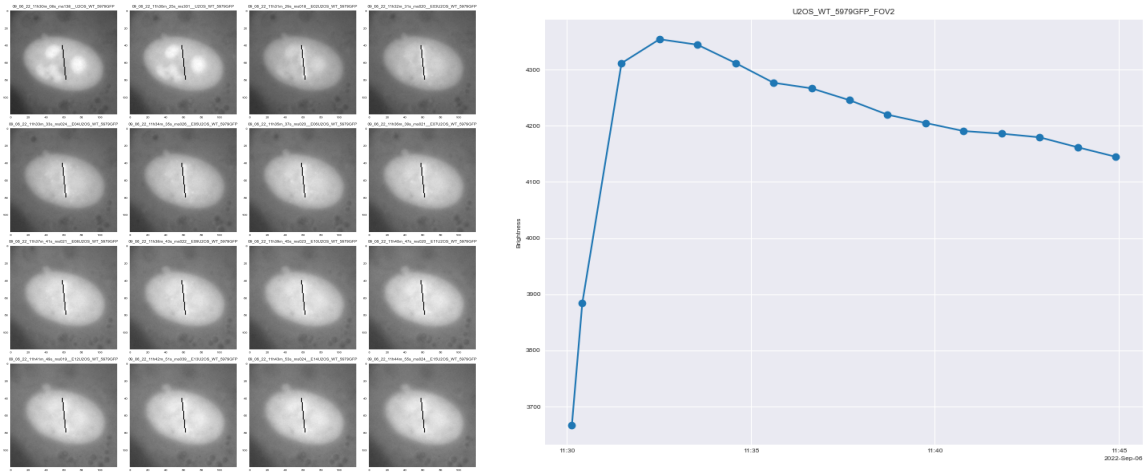


Figure 2(a): Recruitment Le Detection in a Single Cell, **(b):** Quantification of GFP Intensity Over Time

3.2 Automated Image Processing of Calcium Responses After Laser-Induced Shockwave Injury

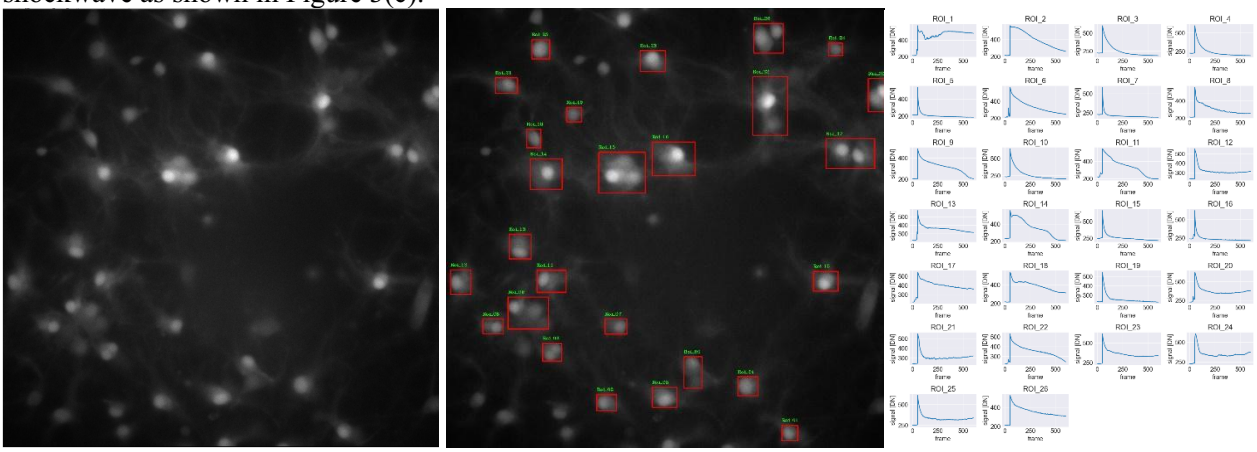
Building upon the expertise developed in automating DNA repair imaging analysis, we extended our machine learning-driven methods to a second biological application: the quantitative analysis of cellular calcium dynamics following LIS. This application represents a natural progression, leveraging the core strengths of automated image segmentation and intensity quantification techniques to solve a new but equally complex biological problem.

LIS systems simulate localized mechanical injuries at the cellular level, enabling the study of intracellular calcium fluxes: a critical early indicator of neuronal and astrocytic responses to TBI conditions. However, manually analyzing the large volumes of time-lapse imaging data generated by these experiments is impractical and susceptible to significant observer bias. Therefore, an efficient, automated workflow was designed to systematically extract, quantify, and visualize calcium response patterns across hundreds of cells over multiple frames.

The workflow we established for LIS data analysis involves several key innovations:

1. **Brightest Frame Indexing**: To accurately detect the initiation of the calcium signaling response, the entire image series was first scanned to identify the frame exhibiting the highest overall fluorescence intensity, marking the onset of the shockwave effect, as shown in Figure 3(a).
2. **Automated ROI (Region of Interest) Segmentation**: Cells were automatically segmented from the frame containing the peak fluorescence signal. Using adapted algorithms initially developed for DNA repair studies, each cell was delineated as an individual Region of Interest (ROI), enabling targeted intensity tracking as shown in Figure 3(b).
3. **Temporal Tracking of Cellular Response**: For each segmented ROI, fluorescence intensity was measured and recorded across all frames, capturing the temporal evolution of the calcium response at the single-cell level.
4. **Batch Processing of Multiple ROIs**: The system efficiently looped through all detected ROIs and frames, automatically compiling time-resolved fluorescence traces for large populations of cells, with minimal manual intervention.
5. **Quantitative Data Output and Visualization**: The intensity trajectories of each cell were exported into structured datasets (CSV format), enabling downstream statistical analysis. Visualization tools such as Matplotlib were

employed to generate comprehensive response curves, facilitating rapid interpretation of cellular behavior post-shockwave as shown in Figure 3(c).



(a) Fluorescence Image **(b)** Automated Cell Detection **(c)** Calcium Responses

Figure 3: Quantitative Tracking of Calcium Responses Across All Segmented Cells Over Time

This pipeline **markedly accelerated** the analysis process, reducing what previously required several hours of manual tracing per dataset to 2-4 minutes of fully automated computation. Moreover, the automated system maintained a high level of accuracy in identifying subtle and heterogeneous calcium responses among different cell types, offering an efficient and highly scalable solution for large experimental datasets.

By transferring the machine learning experience gained from DNA repair imaging to the field of traumatic injury modeling, we demonstrated the **robust adaptability and transformative potential** of AI-enhanced image analysis across distinct yet biologically significant applications.

3.3 Automated Calcium Imaging Analysis in Mouse Primary Cortical Neurons

Building upon the successful machine learning frameworks developed for DNA repair and LIS calcium studies, we extended our methods to investigate calcium flux in mouse primary cortical neurons. Cortical neurons are particularly vulnerable in neurodegenerative diseases such as Alzheimer's disease. Studying cellular calcium dynamics post-LIS provides key information on early functional deficits associated with disease progression.

Traditionally, analyzing neuronal calcium transients relied on MATLAB-based approaches, which were prone to instability, frequent error messages, and extensive manual masking efforts. To overcome these limitations, we designed a robust, fully automated Python pipeline that significantly enhanced analysis speed, reliability, and accessibility.
The workflow we developed consists of the following steps:

1. **Input Image Selection**: Users provided both Fluo-4 fluorescence images (capturing calcium signals) and phase contrast images, along with channels for Dead Red (dead cells) and empty background views.
2. **Dead Cell Detection**: A watershed segmentation algorithm was applied to Dead Red channels to accurately identify non-viable cells, ensuring they were excluded from subsequent calcium response analyses.
3. **Live Cell Segmentation**: Cellpose, a generalist AI-based cell segmentation algorithm, was utilized to automatically delineate individual live neurons (Figure 4), eliminating the need for manual outlining.
4. **Shockwave Frame Detection and Validation**: The system automatically detected the frame at which the LIS-induced calcium response peaked and validated intensity data quality for each cell.
5. **Quantitative Calcium Analysis**: For each valid live cell, $\Delta F/F_0$ values were computed to normalize calcium intensity changes over time. Calcium flux peaks were detected, and two-phase exponential decay models were fitted to estimate transient half-lives, providing detailed kinetic profiles.

6. **Data Export and Visualization**: The pipeline generated time-course plots (e.g., Figure 5) highlighting shockwave onset, calcium peak detection, and return to baseline. The final results were exported in CSV format for easy statistical analysis.

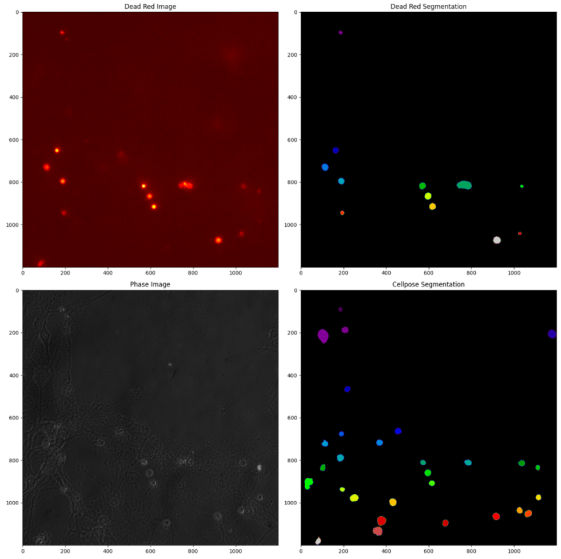


Figure 4: Cellpose & Dead red segmentation

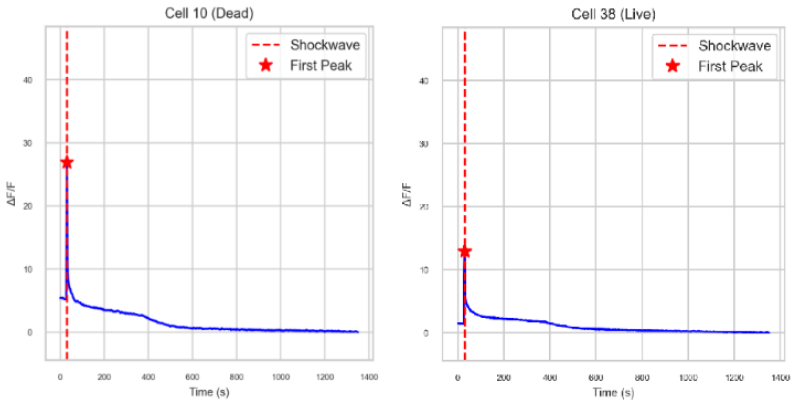


Figure 5: Calcium Response of Dead (left) and live (right) cells

4. Conclusion

In this study, we successfully developed and deployed machine learning-driven image analysis frameworks across three biologically critical applications: DNA repair after laser-induced damage, calcium signaling dynamics following LIS injury, and neuronal calcium flux. By integrating advanced segmentation techniques such as Cellpose with custom Python-based automation, we significantly accelerated data processing, increased analytical accuracy, and eliminated subjective bias associated with traditional manual methods.

For DNA repair studies, our pipeline achieved rapid and precise quantification of protein recruitment kinetics post-damage, setting a new benchmark for efficient and accurate analysis in live-cell imaging. In the context of TBI modeling, our automated calcium analysis pipeline enabled large-scale temporal mapping of astrocyte and neuronal responses following shockwave stimulation, significantly reducing human labor while maintaining high precision. Finally, by extending these tools to mouse primary cortical neuron models, we demonstrated the versatility and scalability of our machine learning approach in deciphering complex and heterogeneous biological responses.

Our collective achievements showcase the transformative potential of artificial intelligence in biomedical imaging and pave the way for future studies requiring rapid, reproducible, and scalable analysis of dynamic biological phenomena. Continued innovation at the intersection of machine learning and bioengineering will undoubtedly accelerate discoveries in cellular injury, neurodegeneration, and DNA damage response mechanisms.

Acknowledgements

This material was based upon work supported by a gift from Beckman Laser Institute Inc. to LS & VGG. Special thanks to the private donors to our UCSD IEM BTC center: Dr. Shu Chien from UCSD Bioengineering, Dr. Lizhu Chen from CorDx Inc., Dr. Xinhua Zheng, David & Leslie Lee for their generous donations.

.

References

[1] Meselson, M., & Stahl, F. W. (1958). "The Replication of DNA in Escherichia coli." *Proceedings of the National Academy of Sciences*, Volume 44(7), 671-682.

[2] Berns, M.W., Rounds, D.E. (1970). "Cell surgery by laser." *Scientific American*, 222(2):98-103.

[3] Berns, M.W., Wright, W.H., Wiegand Steubing, R. (1991). "Laser microbeam as a tool in cell biology." *International Review of Cytology*, 129:1-44.

[4] Botchway, S.W., Reynolds, P., Parker, A.W., O'Neill, P. (2010). "Use of near infrared femtosecond lasers as sub-micron radiation microbeam for cell DNA damage and repair studies." *Mutation Research*, 704(1-3):38-44.

[5] You, Z., Shi, L.Z., Zhu, Q., Wu, P., Zhang, Y.W., Basilio, A., Tonnu, N., Verma, I.M., Berns, M.W., Hunter, T. (2009). "CtIP links DNA double-strand break sensing to resection." *Molecular Cell*, 36(6):954-69.

[6] Wang, H., Shi, L.Z., Wong, C.L., Han, X., Huang, P.Y., Truong, L.N., Zhu, Q., Shao, Z., Chen, D.J., Berns, M.W., Yates, J.R., Chen, L., Wu, X. (2013). "The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR–Mediated Double-Strand Break Repair." *PLoS Genetics*, 9(2): e1003277.

[7] Lu, C., Truong, L.N., Aslanian, A., Shi, L.Z., Li, Y., Hwang, P.Y., Koh, K.H., Hunter, T., Yates, J.R., Berns, M.W., Wu, X. (2012). "The RING Finger Protein RNF8 Ubiquitinates Nbs1 to Promote DNA Double-strand Break Repair by Homologous Recombination." *Journal of Biological Chemistry*, 287(52).

[8] He, J., Shi, L.Z., Truong, L.N., Lu, C.S., Razavian, N., Li, Y., Negrete, A., Shiloach, J., Berns, M.W., Wu, X. (2012). "Rad50 zinc hook is important for the Mre11 complex to bind chromosomal DNA double-stranded breaks." *Journal of Biological Chemistry*, 287(38):31747-56.

[9] Wang, H., Shao, Z., Shi, L.Z., Hwang, P.Y., Truong, L.N., Berns, M.W., Chen, D.J., Wu, X. (2012). "CtIP Protein Dimerization Is Critical for Its Recruitment to Chromosomal DNA Double-stranded Breaks." *Journal of Biological Chemistry*, 287(25):21471-80.

[10] Truong, L.N., Li, Y., Shi, L.Z., Hwang, P.Y., He, J., Wang, H., Razavian, N., Berns, M.W., Wu, X. (2013). "Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step." *Proceedings of the National Academy of Sciences USA*, 110(19):7720-5.

[11] Wang, H., Li, Y., Truong, L.N., Shi, L.Z., Hwang, P.Y., He, J., Do, J., Cho, M.J., Li, H., Negrete, A., Shiloach, J., Berns, M.W., Shen, B., Chen, L., Wu, X. (2014). "CtIP maintains stability at common fragile sites." *Molecular Cell*, 54(6):1012-21.

[12] Voris, H.C. (1950). "Mild head injury (concussion)." *American Journal of Surgery*, 80(6):707-13.

[13] Raghupathi, R., McIntosh, T.K., Smith, D.H. (1995). "Cellular responses to experimental brain injury." *Brain Pathology*, 5(4):437-42.

[14] Chen, H., Constantini, S., Chen, Y. (2015). "A Two-Model Approach to Investigate the Mechanisms Underlying Blast-Induced Traumatic Brain Injury." In: Kobeissy FH, editor. *Brain Neurotrauma: Molecular, Neuropsychological, and Rehabilitation Aspects.* CRC Press.

[15] Ell, C., Wondrazek, F., Frank, F., Hochberger, J., Lux, G., Demling, L. (1986). "Laser-induced shockwave lithotripsy of gallstones." *Endoscopy*, 18(3):95-6.

[16] Li, F., Park, T.H., Sankin, G., Gilchrist, C., Liao, D., Chan, C.U., Mao, Z., Hoffman, B.D., Zhong, P. (2021). "Mechanically induced integrin ligation mediates intracellular calcium signaling with single pulsating cavitation bubbles." *Theranostics*, 11(12):6090-6104.
[17] Gomez-Godinez, V., Morar, V., Carmona, C., Gu, Y., Sung, K., Shi, L.Z., Wu, C., Preece, D., Berns, M.W. (2021). "Laser-Induced Shockwave (LIS) to Study Neuronal Ca^{2+} Responses." *Frontiers in Bioengineering and Biotechnology*, 9:598896.
[18] Shah SB, Li Y, Li S, Hu Q, Wu T, Shi Y, Nguyen T, Ive I, Shi L, Wang H, Wu X. 53BP1 deficiency leads to hyperrecombination using break-induced replication (BIR). Nat Commun. 2024 Oct 5;15(1):8648. doi: 10.1038/s41467-024-52916-z. PMID: 39368985; PMCID: PMC11455893
[19] Pachitariu, M., & Stringer, C. (2022). Cellpose 2.0: how to train your own model. Nature Methods, 19(12), 1634–1641.