A Deep Learning Pipeline for Lipid Droplets Analysis in Glioblastoma Spheroids

Maroa Alani[1], Marie Beurton-Aimar[2], Lucie Brisson[3], and Cloé Tessier[4]

[1] LaBRI, UMR 5800, University of Bordeaux, France
maroa.alkhalaf-alkharaba@u-bordeaux.fr, marie.beurton@labri.fr
[2] BRIC, UMR 1312, Inserm, University of Bordeaux, France
lucie.brisson@inserm.fr,cloe.tessier@u-bordeaux.fr

Abstract. Glioblastoma is an aggressive brain tumor in which altered lipid metabolism is closely linked to treatment resistance. Especially, Lipid Droplets(LDs) accumulate in glioblastoma spheroids and represent an important metabolic biomarker, motivating the need for accurate and reproducible quantification. However, manual LDs analysis in confocal microscopy images remains time consuming, subjective, and difficult to scale. We propose a pipeline for automated LDs segmentation and quantification in 2D confocal microscopy images. This method is based on a U-Net-dc architecture enhanced with dilated convolutions and trained on 640 annotated images using a hybrid Dice/Focalloss. The pipeline integrates background correction, data augmentation, batch level inference, binary mask generation, per image statistics, and aggregated quantitative reports. The U-Net-dc model achieved a Dice score of 0.97, an accuracy of 98.1%, and high precision and recall (approximately 98% and 96%, respectively), showing clear improvement over classical image processing methods. In conclusion, the proposed U-Net-dc pipeline enables robust and scalable high throughput LDs analysis and offers a foundation for future extensions to detection based and real time microscopy applications. Code will be made publicly available at: https://github.com/malani86/unet-DC-segmentation.git

Keywords: Glioblastoma · Lipid Droplets(LDs) · Deep Learning · U-Net-dc · Fluorescence Microscopy · Biomedical Image Analysis · Semantic Segmentation .

1 Introduction

Glioblastoma is the most aggressive primary brain tumor in adults and is characterized by strong intratumoral heterogeneity, rapid progression, and poor response to therapy. Recent studies have highlighted the role of altered lipid metabolism in glioblastoma progression, especially the accumulation of intracellular lipid droplets (LDs), which are associated with metabolic adaptation and treatment resistance [1]. In fluorescence microscopy images of glioblastoma spheroids, LDs appear as small, bright, and heterogeneous structures, making

their quantification important for downstream biological analysis. However, reliable LDs quantification remains difficult in practice. Classical workflows typically rely on illumination correction, thresholding, and watershed based separation of touching droplets. While usable for small scale studies, these methods are sensitive to threshold choice, imaging variability, and object crowding, which reduces reproducibility and makes large scale analysis cumbersome [2, 3]. Deep learning has significantly improved biomedical image segmentation by learning task-specific representations directly from annotated data, and generally outperforms hand crafted pipelines in robustness and generalization [5]. Among existing architectures, U-Net has become a standard solution for pixel wise segmentation in microscopy due to its encoder–decoder design and skip connections [4]. In this work, we investigate a dilated convolution variant, U-Net-dc, to improve the detection of small and heterogeneous droplets in complex fluorescence backgrounds. We present a pipeline for automated LDs segmentation and quantification in 2D fluorescence microscopy images of glioblastoma spheroids, and evaluate it on 640 annotated images for scalable batch level biological analysis.

2 Related Work

Automated analysis of fluorescence microscopy images has traditionally relied on classical image-processing techniques such as thresholding, morphological filtering, and watershed-based separation. In lipid-droplet analysis, these methods are attractive because droplets often appear as bright compact structures against a darker background. However, their performance remains sensitive to threshold selection and to the separation of connected droplets, which reduces robustness in heterogeneous or crowded images. Earlier machine-learning approaches improved flexibility but still relied on handcrafted features and limited local context [13]. Deep learning has since become central to biomedical image segmentation because it learns hierarchical representations directly from annotated data and generally offers stronger robustness than classical methods [5]. Among the available architectures, U-Net has become a standard model due to its encoder–decoder design and skip connections, which preserve fine spatial information while integrating multiscale context [4]. Several extensions have been proposed to improve localization and feature reuse in microscopy tasks [19–21], including dilated-convolution variants that enlarge the receptive field without sacrificing spatial resolution, which is beneficial for small and heterogeneous biological structures [17]. Recent studies have also applied deep learning to lipid-droplet and organelle analysis in microscopy, where segmentation outputs are directly linked to downstream measurements such as object count, size, and spatial distribution [12, 14, 11]. Automated deep-learning pipelines are therefore increasingly used in high-throughput microscopy and subcellular phenotyping [7, 6]. However, fewer studies examine robustness under external or cross-dataset evaluation, even though differences in staining, illumination, acquisition hardware, and morphology can substantially affect model performance [15, 16]. Moti-

vated by these observations, the present work investigates a U-Net based pipeline with dilated convolutions for lipid droplet segmentation and quantification.

3 Methods

3.1 Classical Baseline for Lipid Droplet Segmentation

As a classical reference, we considered the semi automated BlobInspector based workflow developed in the previous stage of the project. This pipeline was specifically designed for lipid droplet analysis in fluorescence microscopy images and combines several conventional image processing steps, including background correction, Gaussian filtering, local thresholding, size filtering, connected component labeling, and watershed separation of touching droplets. In this workflow, lipid droplets are identified as bright compact structures after thresholding, while post processing removes isolated pixels and small spurious detections. Watershed segmentation is then applied to separate adjacent droplets and enable droplet level quantification. The resulting objects are used to compute biologically relevant measurements such as droplet count, area, equivalent diameter, and centroid position. The pipeline also exports structured results in CSV/Excel format for downstream statistical analysis. Although this classical approach is computationally simple and provides interpretable intermediate steps, it requires manual parameter tuning and remains sensitive to illumination artifacts, image noise, and object crowding. In particular, faint droplets may be missed, while touching droplets may be merged or incorrectly separated depending on the thresholding and watershed settings. These limitations motivated the transition toward a deep learning based segmentation framework.

3.2 Deep learning model

We employed a U-Net based semantic segmentation network enhanced with dilated convolutions,introduced by Ji et al. (2023)[17],and referred to as U-Net-dc. As in the original U-Net architecture, the model follows an encoder–decoder structure with skip connections between corresponding stages of the contracting and expanding paths [4]. The encoder extracts hierarchical contextual features, while the decoder reconstructs spatial details to produce a dense segmentation map. Unlike the baseline U-Net, U-Net-dc architecture as shown in figure.1[17] incorporates dilated convolutions in the deeper layers of the network. Dilated convolutions increase the receptive field without increasing the number of trainable parameters or reducing feature map resolution, making them suitable for detecting small and heterogeneous Lipid Droplets in fluorescence microscopy images. Previous studies have shown that dilated U-Net variants can improve contextual feature extraction while preserving fine spatial details in biomedical image segmentation tasks [17]. The network outputs a single-channel probability map representing the likelihood of each pixel belonging to the lipid droplet class.

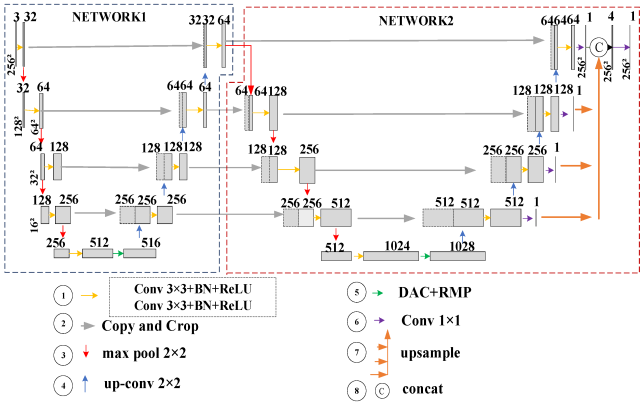


Fig. 1: U-Net-dc Architecture (U-Net with Dilated Convolutions)

3.3 Loss Functions Strategy and Model optimization

Because Lipid Droplets occupy only a small fraction of the image area, the segmentation task is strongly affected by class imbalance. To address this issue, the proposed model was trained with a hybrid loss combining Dice loss [9] and binary Focal loss [8]. Let $p_i \in [0,1]$ denote the predicted probability for pixel i, and let $g_i \in \{0,1\}$ denote the corresponding ground truth label. Dice loss is defined as:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2\sum_i p_i g_i + \epsilon}{\sum_i p_i + \sum_i g_i + \epsilon}, \tag{1}$$

where ϵ is a small constant added for numerical stability. Binary Focal loss is defined as:

$$\mathcal{L}_{\text{Focal}} = -\frac{1}{N}\sum_i \alpha\, g_i (1-p_i)^{\gamma} \log(p_i) - \frac{1}{N}\sum_i (1-g_i)\, p_i^{\gamma} \log(1-p_i), \tag{2}$$

where α controls class weighting, γ is the focusing parameter, and N is the total number of pixels. The final training objective is defined as a weighted combination of the two losses:

$$\mathcal{L}_{\text{total}} = 0.7\,\mathcal{L}_{\text{Dice}} + 0.3\,\mathcal{L}_{\text{Focal}}. \tag{3}$$

In our implementation, the focal component used $\alpha = 1.0$ and $\gamma = 2.0$. This hybrid objective was selected to improve overlap quality while maintaining robustness to severe foreground/background imbalance. Model optimization was performed with Adam using an initial learning rate of 10^{-3}. A ReduceLROnPlateau scheduler was used to reduce the learning rate by a factor of 0.5 when validation performance plateaued for 5 epochs. Training used a batch size of 8 and ran for up to 15 epochs, with early stopping based on validation Dice score and a patience of 5 epochs .

3.4 Creating Ground Truth

Our study is based on fluorescence microscopy images of glioblastoma spheroids acquired from confocal microscopy observations of in vitro cell spheroids. Lipid Droplets (LDs) were visualized after fluorescent labeling, and the original microscopy data consisted of grayscale 16-bit images with spatial dimensions of 1024×1024 pixels and intensity values ranging from 0 to 4095. In previous work

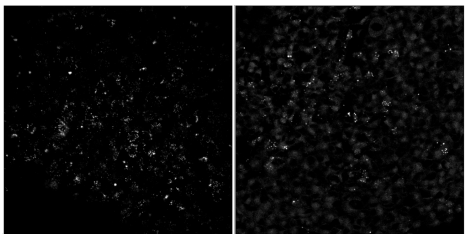


Fig. 2: Glioblastoma Lipid Droplets accumulation and examples of microscopy images used in this study.

we have developed a software, BlobInspector, designed to analyze blobs in biological images, the code will be made publicly available at: https://github.com/lbssn/BlobInspector.git. It provides several common tools for computing metrics related to the blobs, such as coordinates, size, density, and distances. In BlobInspector, segmentation is implemented with classical algorithms. Where combines rolling-ball background correction, intensity thresholding, a new binary image is obtained. The white pixels are set to figure out pixels belonging to the droplet. Finally, isolated pixels are suppressed by post treatment. In BlobInspector, we label droplets by identifying individual droplets obtained by applying the watershed algorithm. This operation allows to quantify how many droplets are present on the images. From BlobInspector, we have been able to semi automatise the images labeling and create a kind of ground truth data set. From left to right Fig.3 : (1) the

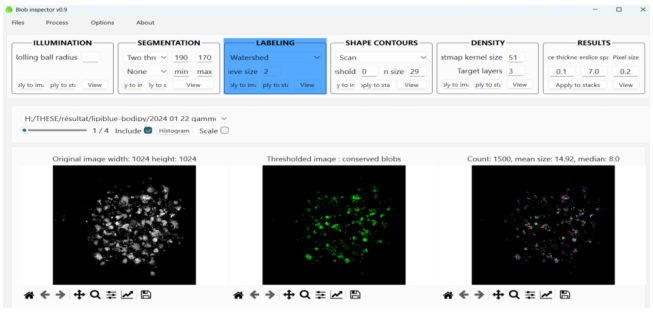


Fig. 3: Classical BlobInspector-based Lipid Droplets analysis workflow.

original fluorescence microscopy image of a glioblastoma spheroid (1024 × 1024 pixels), showing heterogeneous intracellular lipid-droplet accumulation; (2) the threshold image obtained after rolling-ball background correction and intensity based blob preservation, where candidate lipid droplet regions are highlighted in green; and (3) the connected component analysis output, where individual detected droplets are assigned different colors for visualization and quantitative analysis. Although the workflow enables semi automated droplet quantification, segmentation quality remains sensitive to threshold selection and droplet crowding, particularly in dense spheroid regions where neighboring droplets may merge or fragment.

3.5 Evaluation Metrics

Performance was assessed on separate test set using pixel accuracy, Dice score, precision, recall, F1-score, specificity, and confusion-matrix analysis. Precision, recall, F1-score, and specificity were computed after binarizing predicted probability maps with a threshold of 0.3, consistent with the inference pipeline used during validation and testing . In addition to quantitative evaluation, qualitative performance was further examined using predicted masks, difference maps, and overlay visualizations on representative test images .

4 Results

4.1 Experimental Setup

Lipid Droplets Dataset A supervised dataset of 640 annotated 2D images was constructed by using semi automated labels derived from the earlier BlobInspector workflow. Pairs of images and there masks were divided using a two stage random split implemented in the training script. First, 20% of the dataset was reserved for testing. The remaining 80% was then split into 75% training and 25% validation,resulting in a final split of 60% training, 20% validation, and 20% testing .

Preprocessing and Data Augmentation Each image was loaded in RGB format and underwent rolling-ball background correction before training and inference. In the implemented data pipeline, background correction was applied independently to each channel using a morphological opening with a rolling-ball radius of 50 pixels which determined experimentally, followed by channel wise subtraction and normalization. After correction, images and masks were resized to a fixed spatial resolution of 512 × 512 pixels. Image intensities were normalized to the range $[0, 1]$, and masks were converted into binary masks with background and foreground values of 0 and 1, respectively. To reduce overfitting and improve model generalization, online data augmentation was applied during training. The augmentation pipeline included horizontal flips, vertical flips,

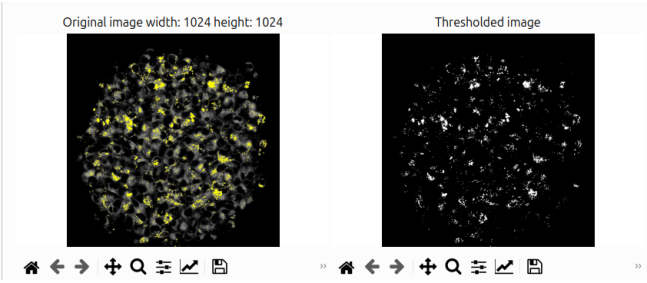


Fig. 4: Example of semi-automated ground-truth generation for Lipid Droplets segmentation.

random 90 degree rotations, brightness and contrast perturbations, and elastic deformations.These augmentations were introduced to improve robustness to intensity and morphological variability.

4.2 Performance Evaluation

Table 1 summarizes the segmentation performances obtained on the test set. U-Net-dc model achieved a Dice score of 0.97 and a pixel accuracy of 98.1%, together with high precision and recall, indicating strong agreement between predicted masks and reference annotations.

Metric	Score
Accuracy	**98.1%**
Dice score	**0.97**
Precision	**98.2%**
Recall	**96.0%**

Table 1: Segmentation performance

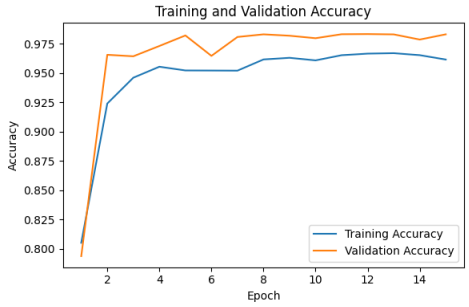


Fig. 5: Accuracy plot of the U-Net-dc model.

These results indicate that the network correctly recovers most droplet pixels while maintaining a low false positive rate. The high precision suggests that predicted droplet regions are highly reliable, whereas the strong recall shows that the majority of annotated droplets are successfully detected. Overall,the model generates reliable segmentation masks for quantitative analysis.

Learning Curves and Confusion Matrix Training and validation curves showed stable convergence over the optimization process.the model converged within 15 epochs while maintaining similar training and validation performance and no major divergence between the two curves, indicating that overfitting remained limited during training. Likewise, the joint evolution of loss and Dice score indicated progressive improvement during optimization and stable validation behavior in later epochs. The observed training stability is likely related to

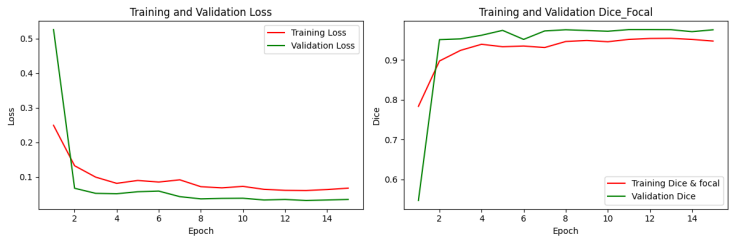


Fig. 6: loss and dice focal plot U-Net-dc

the adopted training strategy, which combines a hybrid Dice/Focal loss, online data augmentation, learning rate scheduling, and early stopping. These strategies helped stabilize training under severe class imbalance. The confusion matrix computed on the test set further confirms the robustness of the proposed model. The test predictions exhibit high sensitivity and specificity for the droplet class, with relatively limited confusion between foreground and background pixels. This behavior is especially important in the present setting, where the foreground class occupies only a small fraction of the image and where even a modest number of false positives can noticeably affect droplet quantification. From the values in

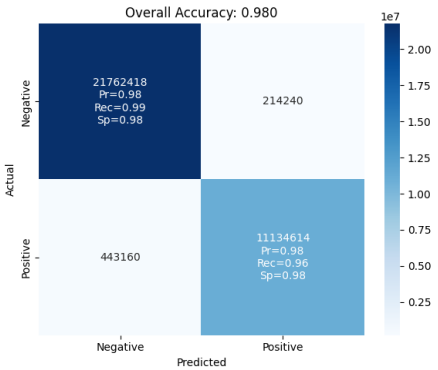


Fig. 7: confusion matrices U-Net-dc

confusion matrix (figure 7), we calculated the following metrics for the droplet class:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad \text{Specificity} = \frac{TN}{TN + FP}.$$

true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Comparison with U-Net To better assess the contribution of the proposed dilated convolution design, we compared U-Net-dc with a standard U-Net baseline trained on the same lipid droplet dataset under the same experimental setting. Table 2 summarizes the obtained results. U-Net-dc achieved higher overall accu-

Table 2: Comparison of segmentation methods on the main lipid droplet dataset.

Method	Accuracy (%)	Dice	Precision (%)	Recall (%)
U-Net	97.0	–	94.0	97.0
U-Net-dc	98.1	0.97	98.2	96.0

racy than the standard U-Net (98.1% vs. 97.0%). Although the standard U-Net obtained slightly higher recall, U-Net-dc provided higher precision, indicating fewer false positives and more reliable lipid droplet quantification.

4.3 Inference and Quantification Pipeline

During inference, unseen images were processed using the same main preprocessing steps as during training, namely RGB conversion, rolling-ball correction, resizing to 512×512, and intensity normalization. The trained network produced a probability map, which was binarized using a threshold of 0.3 . Small detections below 2 pixel^2 were removed during post processing in order to suppress isolated artifacts . The resulting binary masks were then analyzed to extract quantitative droplet measurements. Connected component analysis identified individual droplets and measured their area, diameter, centroid position, and total count.The pipeline also generated per image summary tables and exported aggregated measurements in CSV or spreadsheet format for comparison across experimental conditions. The implementation of the proposed pipeline is available at: https://github.com/malani86/unet-DC-segmentation.git.

4.4 Visual Analysis

Overall, the confusion matrix analysis is consistent with the quantitative results reported in Table 1: most foreground pixels are correctly identified, and the model maintains clear separation between LDs regions and the surrounding

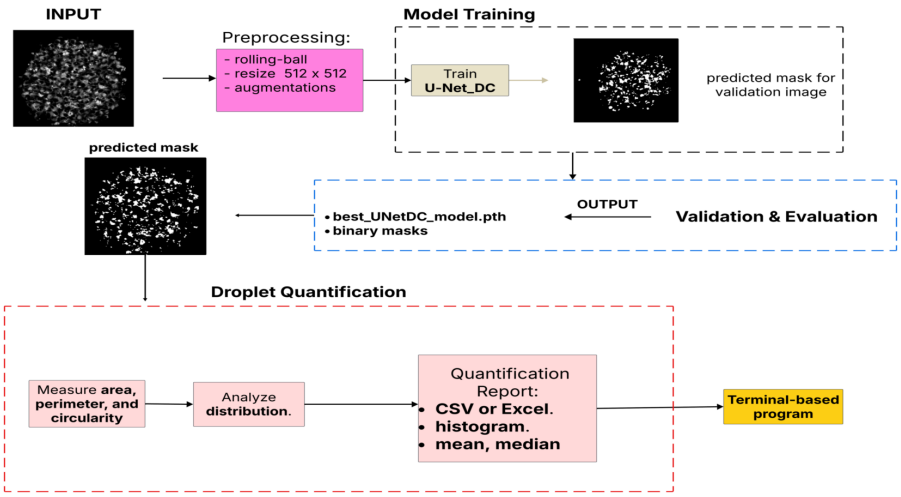


Fig. 8: U-Net-DC segmentation pipeline

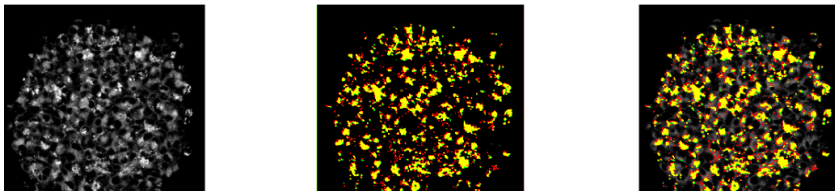


Fig. 9: Segmentation results showing the predicted mask and overlay for image 17 (complex scene).

background. We overlaid the predicted masks on the original images and also generated difference maps to visualize errors. The overlay displays the model's predicted droplet regions in a translucent color on top of the raw image, allowing visual comparison with the ground truth mask. The difference map highlights false positives (FP) and false negatives (FN), for instance, we marked false Negatives (droplet pixels missed by the model) in red and false Positives (pixels predicted as droplet but not in the ground truth) in green. Visual inspection of selected test images shows that the proposed approach accurately segments LDs across a range of image conditions, including both sparse and crowded droplet distributions. The examples shown demonstrate that the predicted masks generally follow the annotated droplet regions closely, while maintaining consistent segmentation boundaries. The saved overlay and difference map outputs from the testing script the agreement between predictions and reference masks. Most discrepancies are localized around droplet boundaries or within highly clustered regions, which is consistent with the expected difficulty of separating small adjacent objects in fluorescence microscopy images.

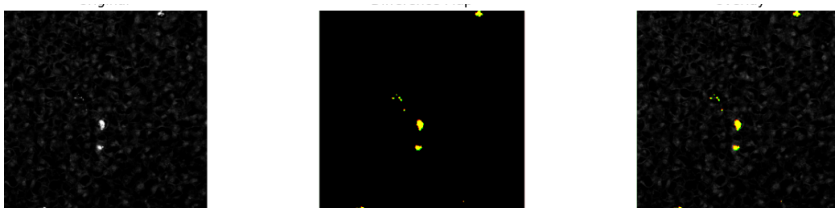


Fig. 10: Segmentation results showing the predicted mask and overlay for image 153 (sparse scene).

Table 3: Comparison of BlobInspector and the proposed deep learning pipeline on two representative lipid droplet images.

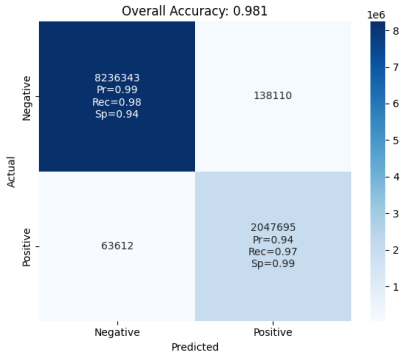
Image	Method	Droplet count	Total area (px)
image153	BlobInspector	10	90
image153	Deep learning	10	542
image17	BlobInspector	291	1400
image17	Deep learning	351	7703

Comparison with the Classical Quantification Workflow To complement the segmentation based evaluation, we compared the proposed deep learning pipeline with the earlier BlobInspector classical workflow on two representative lipid droplet images, including a sparse example (image153) and a dense example (image17). For each method, we measured the number of connected components and the total segmented area obtained from the corresponding binary masks. Table 3 shows that the deep learning pipeline generally produced larger segmented areas than BlobInspector and detected more droplets in the dense image. This suggests higher sensitivity to weak or clustered lipid droplets and highlights the effect of segmentation quality on downstream quantification. In particular, it yields larger segmented areas and, in dense regions, higher droplet counts, which is consistent with improved sensitivity to small or faint lipid droplets.

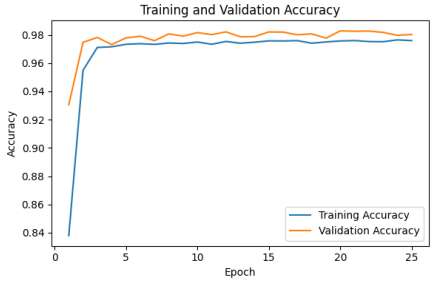
4.5 External Evaluation on the BBBC039 Dataset

To assess whether the proposed U-Net-dc architecture could be applied beyond lipid droplet segmentation, additional experiments were conducted on the public BBBC039 fluorescence nucleus segmentation dataset [10]. The BBBC039 dataset contains 200 fluorescence microscopy images of cell nuclei with expert provided annotations. For this study, the dataset was divided into 60% training, 20% validation, and 20% testing subsets, corresponding to 120 training images, 40 validation images, and 40 test images.To match the binary segmentation setting used in this study, the original annotations were converted into foreground/background masks. The exact same U-Net-dc architecture used in the lipid droplet experiments was retrained on BBBC039 using the same implementation and training pipeline as in the main study. The model produced a binary segmentation mask

for each input image, and performance was assessed using quantitative metrics, confusion matrix analysis, and qualitative visual inspection through overlay images and difference maps. Quantitatively, the retrained model achieved an overall



(a) Confusion matrix analysis of the U-Net-dc model on the BBBC039 dataset.



(b) Training and validation accuracy curves obtained during training on the BBBC039 dataset.

Fig. 11: Performance evaluation of the U-Net-dc model on the BBBC039 dataset.

accuracy of 98.1%. The confusion matrix showed that most foreground and background pixels were correctly classified, with precision of approximately 94% and recall of approximately 97% for the nucleus class. Despite the strong class imbalance toward background pixels, the model maintained good separation between nucleus and background regions. The training and validation curves also showed rapid performance improvement during the first epochs, followed by stable convergence. Although moderate fluctuations were observed in the validation loss, the validation accuracy and Dice score remained consistently high, with no clear evidence of severe overfitting. Qualitative evaluation confirmed these results: pre-

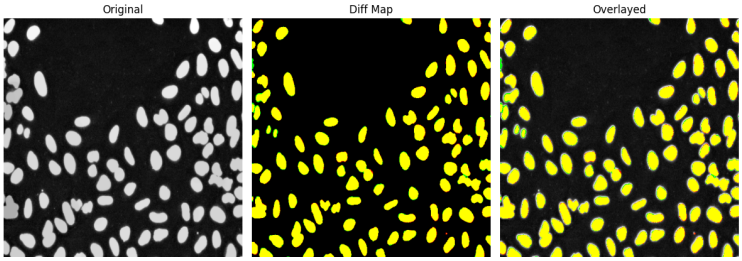


Fig. 12: Qualitative evaluation of nucleus segmentation on the BBBC039 dataset.

dicted masks aligned well with reference annotations, with most nuclei correctly

localized, including closely adjacent instances. Remaining errors were concentrated at boundaries and in touching or ambiguous nuclei, consistent with known challenges in nucleus segmentation [18]. Compared to U-Net based nucleus segmentation benchmarks [18], the proposed U-Net-dc achieved competitive results despite its simpler binary formulation, suggesting that dilated convolutions help capture broader context while preserving spatial detail in densely packed structures. Overall, these findings indicate that U-Net-dc generalizes beyond lipid droplet segmentation and may be applicable to related fluorescence microscopy tasks after appropriate retraining.

5 Discussion

The proposed U-Net-dc pipeline achieved strong segmentation performance, with a Dice score of 0.97 and accuracy of 98.1%, indicating good agreement with reference annotations while keeping false detections limited. Beyond producing masks, the pipeline extracts droplet level and image level measurements, including count, area, and size statistics, making it suitable for high-throughput biological studies where manual analysis would not scale [7, 6]. Compared to the classical thresholding and watershed workflow, the learned approach behaved more consistently across varying image conditions, particularly in regions with uneven contrast or high droplet density where classical methods often struggled with merging or over segmentation. Similar improvements have been reported in other fluorescence microscopy studies [18, 22]. Several limitations remain. First, the analysis is restricted to 2D, whereas glioblastoma spheroids are three dimensional structures, extending the pipeline to volumetric data would provide a more complete characterization of LD organization. Second, training annotations were derived from a semi-automated workflow and may contain residual inconsistencies that could affect evaluation. Third, the framework performs semantic rather than instance segmentation, so separation of touching droplets still relies on post processing, future work could explore detection based or instance segmentation approaches such as YOLO or Mask R-CNN for more direct droplet counting in crowded scenes. Additional directions include cross dataset validation, adaptation to 3D stacks, and real time microscopy integration.

6 Conclusion

Overall, the results suggest that the proposed U-Net-dc pipeline can serve as a practical tool for automated Lipid Droplets analysis in glioblastoma microscopy images. In our experiments, the method produced stable segmentations across different image conditions while reducing the amount of manual adjustment required by the earlier workflow. The combination of automated segmentation and quantitative reporting also makes the pipeline suitable for larger scale microscopy studies. More broadly, these results support the use of deep learning based methods for quantitative analysis of altered lipid metabolism in glioblastoma spheroid models.

References

1. Wu, X., Geng, F., Cheng, X., Guo, Q., Zhong, Y., Cloughesy, T.F., Yong, W.H., Chakravarti, A., Guo, D.: Lipid droplets maintain energy homeostasis and glioblastoma growth via autophagic release of stored fatty acids. *iScience* **23**(10), 101569 (2020). https://doi.org/10.1016/j.isci.2020.101569
2. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* **9**(1), 62–66 (1979). https://doi.org/10.1109/TSMC.1979.4310076
3. Beucher, S.: The watershed transformation applied to image segmentation. *Scanning Microscopy Supplement* **1992**(6), Article 28 (1992).
4. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-319-24574-4_28
5. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data* **8**, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8
6. Melanthota, S.K., Gopal, D., Chakrabarti, S., Kashyap, A.A., Radhakrishnan, R., Mazumder, N.: Deep learning-based image processing in optical microscopy. *Biophysical Reviews* **14**, 463–481 (2022). https://doi.org/10.1007/s12551-022-00949-3
7. Bilodeau, A., et al.: Automated microscopy image segmentation and analysis with machine learning. In: *Fluorescent Microscopy*. Methods in Molecular Biology. Springer (2022). https://doi.org/10.1007/978-1-0716-2051-9_20
8. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988 (2017). https://doi.org/10.1109/ICCV.2017.324
9. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M.J.: Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 240–248. Springer (2017). https://doi.org/10.1007/978-3-319-67558-9_28
10. Broad Bioimage Benchmark Collection: BBBC039v1, nuclei of U2OS cells in a chemical screen. Available at: https://bbbc.broadinstitute.org/BBBC039. Last accessed 12 May 2026.
11. Egebjerg, J.M., Szomek, M., Thaysen, K., Juhl, A.D., Kozakijevic, S., Werner, S., Pratsch, C., Schneider, G., Kapishnikov, S., Ekman, A., Röttger, R., Wüstner, D.: Automated quantification of vacuole fusion and lipophagy in *Saccharomyces cerevisiae* from fluorescence and cryo-soft X-ray microscopy data using deep learning. *Autophagy* **20**(4), 902–922 (2024). https://doi.org/10.1080/15548627.2023.2270378
12. Jena, L., Shanthi, S., Devi, A.G., Nayak, S.R., Biswas, P.: Lipid droplet segmentation using U-Net convolutional neural network architecture. In: *AIP Conference Proceedings* **3122**, 030021 (2024). https://doi.org/10.1063/5.0215943
13. Dejgaard, S.Y., Presley, J.F.: New automated single-cell technique for segmentation and quantitation of lipid droplets. *Journal of Histochemistry & Cytochemistry* **62**(12), 889–901 (2014). https://doi.org/10.1369/0022155414554835
14. Liu, Z., Zhang, H., Jin, L., Chen, J., Nedzved, A., Ablameyko, S., Ma, Q., Yu, J., Xu, Y.: U-Net-based deep learning for tracking and quantitative analysis of intracellular vesicles in time-lapse microscopy images. https://doi.org/10.1142/S1793545822500316

15. Xu, S., et al.: GeNSeg-Net: A general segmentation framework for any nucleus in immunohistochemistry images. In: *Proceedings of the 32nd ACM International Conference on Multimedia* (2024). https://doi.org/10.1145/3664647.3681441
16. Chen, S., Ding, C., Liu, M., Cheng, J., Tao, D.: CPP-Net: Context-aware polygon proposal network for nucleus segmentation. *IEEE Transactions on Image Processing* **32**, 980–994 (2023). https://doi.org/10.1109/TIP.2023.3237013
17. Ji, Z.; Yao, D.; Chen, R.; Lyu, T.; Liao, Q.; Zhao, L.; Ganchev, I. U-Net_dc: A Novel U-Net-Based Model for Endometrial Cancer Cell Image Segmentation. *Information* **14**(7), 366 (2023). https://doi.org/10.3390/info14070366
18. Caicedo, J.C.; Roth, J.; Goodman, A.; Becker, T.; Karhohs, K.W.; Broisin, M.; Molnar, C.; McQuin, C.; Singh, S.; Theis, F.J.; Carpenter, A.E. Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A* **95**(9), 952–965 (2019). https://doi.org/10.1002/cyto.a.23863
19. Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.: Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999* (2018). https://doi.org/10.48550/arXiv.1804.03999
20. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: A nested U-Net architecture for medical image segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1
21. Long, F.: Microscopy cell nuclei segmentation with enhanced U-Net. *BMC Bioinformatics* **21**, 8 (2020). https://doi.org/10.1186/s12859-019-3332-1
22. Ho, D.J., Fu, C., Salama, P., Dunn, K.W., Delp, E.J.: Nuclei segmentation of fluorescence microscopy images using three dimensional convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 834–842 (2017).