**Keywords:** Intestinal Stem Cells, Cells Detection, Cells Tracking, Deep Learning

**Abstract**

Studying homeostasis stem cells holds the potential to unlock an understanding of the intricate stem cell's proliferation and differentiation. Deciphering the movements of fluorescent kinase translocation reporters (KTRs) in intestinal stem cells (ISCs) can provide valuable insights into the time-dependent and spatial dynamics within signalling pathways, particularly those determining cell fate within the extracellular signal-regulated kinase (ERK) node of the Ras/MAPK pathway. The bioimage analysis is a crucial aspect of investigations, and image segmentation serves as its foundation. Segmenting kinase translocation reporter's (KTRs) microscopy images manually is time-consuming, and inimitable. Although FIJI's automated ImageJ Macro scripts offer a computerised segmentation option, it still suffers from extended processing times and potential inaccuracies. This research presents a novel approach using deep learning algorithms to ensure reliable and accurate cell detection and tracking of intestinal stem cell (ISCs) trajectories in video frames captured at different time intervals. Subsequent testing models on testing datasets aimed to bolster model adaptability. The findings demonstrated that when trained on microscopic imaging datasets, object identification algorithms may not invariably detect and trace every cell in a frame. However, the mAP50 value of the YOLOv8 model reached an impressive 0.9073, accompanied by a Multiple Object Tracking Precision of 78.47\%.These results emphasise the potential of using deep learning approaches for studying temporal kinase dynamics in ISCs and speeding the process for future research in stem cell homeostasis bioimage analysis. This is the first attempt to explore the development of deep learning models and tracking algorithms in this context, providing a foundation for future work in this field.

**Introduction**

Digestion, a vital physiological process, is crucial for maintaining good health. The efficient functioning of the epithelial tissues in the intestines is essential for ensuring effective digestion. Intestinal homeostasis and regeneration rely heavily on intestinal stem cells (ISCs), which play a vital role among the various cell types of present. ISCs are controlled by a specialised tissue microenvironment called niche [1] that enables them to serve as the main progenitor cells of the intestine. Given the constant exposure of the intestinal lining to copious physical and chemical stimuli, wear and tear are inevitable. The intestinal crypts contain stem cells that undergo constant division. This process produces progenitor cells, which then develop into specialised epithelial cells. As these cells move up the villi, they go through maturation and eventually shed into the lumen. The process of cycling in intestinal stem cells (ISCs) is tightly regulated and complex. It ensures the continuous renewal of the intestinal epithelium, which is crucial for maintaining gut health and functionality. The examination of ISCs, specifically their intrinsic abilities for growth and specialisation, holds great importance. Maintaining a balance between stem cell quiescence and activity is a defining characteristic of a functioning niche since excessive stem cell proliferation can lead to cancer [2].

The fate determination of intestinal stem cells is greatly influenced by the signalling dynamics, particularly by the extracellular signal-regulated kinase (ERK) node of the Ras/MAPK pathway. To maintain the balance of the intestines, this pathway regulates important cellular functions like proliferation, differentiation, and apoptosis. Disruption of these regulations can lead to under-proliferation, resulting in the loss of tissue integrity, or over-proliferation, which can lead to the development of tumorigenic growth (the development of a tumour in the body)[3]. Understanding how these signalling processes work requires the application of advanced methods such as fluorescent kinase translocation reporters (KTRs). These innovative techniques play a crucial role in studying the dynamics of cellular communication. KTRs provide a distinct benefit by enabling the synchronous monitoring and quantification of signal dynamics in cellular viability. Using these indicators, researchers can acquire valuable knowledge about the spatial and temporal aspects of ERK activation. This, in turn, enhances comprehension of how stem cells make fate determinations within the intestinal epithelium.

The fundamental aspect of the overwhelming majority of extant biological image analysis correlates to the pivotal undertaking at its core of image segmentation, which has a crucial role in outlining and categorising different objects or areas within an image. These tasks are extremely important, particularly when the objective is to gather significant data from intricate biological samples, such as identifying the existence, position, and motion of cell markers. KTRs are manually segmented using FIJI's freehand tool, however, this lacks repeatability and becomes time-consuming for big datasets, making it appropriate for restricted study. Fiji ImageJ Macro scripts that are automated confront restrictions and arduous scripting tasks, resulting in imperfect segmentation outputs.

Arising from these problems, a discernible need for a more sophisticated and effective methodology became evident. Deep Learning (DL), particularly Deep Neural Networks (DNN), has gained significant recognition for its proven effectiveness in many computational visual tasks and has emerged as a promising source of optimism. This study presents a new methodology that utilises the advanced functionalities of a state-of-the-art (SOTA) Deep Neural Network (DNN) object identification model called YOLOv8 (You Look Only Once - Version 8) \cite{yolov8\_ultralytics}. Specifically trained for this purpose, the YOLOv8 model has undergone training to identify intestinal stem cells in video frames obtained through confocal microscopy. To enhance its usefulness post detections, the intestinal stem cells that have been detected are tracked using the ByteTrack algorithm \cite{zhang2022bytetrack}. The overarching aim is to utilise deep learning to achieve precise and recurring detection by subsequently tracking the ISCs. This approach provides a comprehensive solution that includes depictions of bounding boxes, annotating individual cells, and illustrating the paths taken by annotated intestinal stem cells.\par

The first findings derived from this methodology are promising. The YOLOv8-M model demonstrated superior performance compared to other models (YOLOv8-S, YOLOv8-L, YOLONAS-S, YOLONAS-M, and YOLONAS-L) that are trained on the same dataset, as seen by its mAP50 value of 0.9073, mAP50-95 value of 0.4381, and F1 value of 0.81. The ByteTrack algorithm demonstrated exceptional testing accuracy, as seen by its Mean IoU value of 0.7922, MOTP value of 78.47%, and MOTA value of 76.63%. The fundamental problem lies in the constraint of the model, which fails to identify all cells. This limitation results in inefficiencies throughout the monitoring process and interruptions in the continuous tracking of cells. The use of deep neural networks (DNNs) in this particular domain suggests potential for improvement, especially in terms of enhancing accuracy and precision. The investigation of alternate tracking algorithms may provide improved outcomes. However, this endeavour represents a noteworthy advancement in the field, establishing a fundamental structure for the advancement of deep learning models that are specially designed to detect and monitor ISCs.

**Related Work**

Considering related work within the context of Intestinal Stem Cell Signalling Dynamics Analysis (Section 2.1) and Deep Learning in Bio-imaging Analysis (Section 2.2).

2.1 Intestinal Stem Cell Signalling Dynamics Analysis

Extensive study has been conducted to understand the dynamics of stem cell environments, primarily emphasising the crucial signalling mechanisms that govern these niches. Significantly, during the study of Drosophila, it was found that temporal signalling dynamics in maintaining the stem cell niche are crucial \cite{johnson2019iscs}. Despite its importance, the field of Temporal Signal Dynamics is still relatively understudied in the field of cell biology, even though it has a significant influence on the determination of cell fate.\par

Research indicates that the investigation of protein activation or upregulation dynamics has heavily relied on techniques such as Western blotting and enzymatic assays. These methods, conducted at different time intervals, provide valuable information about cellular behaviours by offering a general overview of protein activity in the sampled cell population \cite{greenwald2018fbiosensor}. However, these techniques have shown some constraints mainly in effectively assessing the intricate details of the average cell population \cite{purvis2013iscs}. Fluorescent protein (FP) biosensors have been recognised as more effective alternatives to conventional Western blotting methods, providing valuable information on cellular signalling pathways at the level of individual cells with high temporal resolution in the nanosecond range \cite{purvis2013iscs}. The fluorescence resonance energy transfer (FRET) FP biosensor technique has been recognised as a highly effective method for investigating the temporal dynamics of molecular processes, offering a temporal milliseconds resolution \cite{vilardaga2003gprotein}. Nevertheless, there are concerns regarding photobleaching, and the functionality of this technique requires the utilisation of two different fluorophores, each with distinct spectral characteristics \cite{Kudo2017LivecellMO}. These techniques are mostly investigated inside in-vitro environments. Kinase translocation reporters (KTRs) are extensively employed in the field of kinase dynamics research. These reporters utilise fluorescence to effectively communicate the activity of kinases by indicating their spatial localization between the nucleus and cytoplasm. The Drosophila in-vivo environment serves as a notable model for studying kinase dynamics, namely via the use of kinase translocation reporters (KTRs) and live imaging methods, with a particular focus on the gut \cite{martin2018liveimage}. In the context of evaluating whole-cell kinase activity, KTRs exhibit a higher level of superiority compared to KRET biosensors. The researchers are able to gain an edge in simultaneously multiplexing and monitoring several kinases inside a single cell due to their capacity to detect a single kinase utilising a single fluorophore \cite{Kudo2017LivecellMO}. However, the reliance on an intact nuclear envelope continues to be a constraint, thus limiting their use in situations after nuclear envelope disintegration. \par

Considerable research has been dedicated to examining the activity of ERK (extracellular signal-regulated kinase) in intestinal stem cells (ISCs) through various imaging methods that utilise kinase translocation reporters (KTRs). One significant aspect of analysing biological images, particularly those involving kinase-targeted receptors (KTRs), involves accurately identifying and differentiating between cellular structures such as the nucleus and cytoplasm \cite{timothy2010NF-kB}. The freehand approach, supported by the readily accessible tool FIJI, is commonly used for manual segmentation. However, this technique presents challenges in terms of reproducibility, especially when dealing with large datasets. As a result, its practicality is limited to specific research scenarios \cite{yuenalice2022ktr}. Automation of the segmentation process can be accomplished by using ImageJ Macro scripts within the FIJI software, which serves as a viable alternative. This method entails applying mathematical operations. Notably, developing these scripts consumes a significant amount of time and may lead to inaccurate segmentation results. Thus, manual selection of correctly segmented cells often becomes necessary \cite{valen2016livecell}.

2.2 Deep Learning in Bio-imaging Analysis

Deep convolutional neural networks (ConvNets), which are a type of supervised machine learning, have become increasingly popular in the field of image classification. These neural networks demonstrate exceptional performance via the process of training on annotated datasets, which allows them to accurately assign suitable labels to pictures that they have not seen before \cite{lecun2015deeplearning}\cite{pereira2012convnet}. The study insights into the efficacy and implications of using Convolutional Neural Networks (ConvNets) for the categorization of pixels in pictures \cite{sermanet2014overfeat}\cite{long2015fully}. Convolutional neural networks have shown considerable efficacy in effectively solving real-world scenarios. However, the integration of these methods into the field of biological data processing is still in its early stages. An optimised convolutional neural network was developed to measure fluorescent protein localisation kinase translocation reporters (KTRs), with the aim of providing a generic and shareable solution across multiple labs \cite{valen2016livecell}. The current methodology places emphasis on image segmentation across five different cell lines. This method concentrates on segmenting images from 5 cell lines, while further improvement lies in enhancing the architecture of CNN \cite{valen2016livecell}. The effectiveness of live cell imaging has only been emphasised in recent times. Notably, Ronnenberg et al.'s pioneering work in 2015 won the ISBI cell tracking challenge (CTC) and manifested the prowess of U-Net Convolutional Neural Networks (ConvNets) in accurately segmenting cells in high-resolution microscope images \cite{ronneberger2015unet} \cite{maska202310years}. However, it is worth noting that the optimal solutions for this task remain specific to each dataset, as the datasets themselves exhibit intricate and diverse characteristics. The current methodologies encounter difficulties when analysing videos that have a poor signal-to-noise ratio or when tracking cells that possess intricate forms or patterns \cite{maska202310years}. Deep learning techniques, which are often used in practical applications involving visual tasks, can encounter the issue of overfitting \cite{li2019overfitting}. Due to the inherent characteristics of cell objects, which are non-rigid and non-significant, it is crucial for cell tracking systems to exhibit strong generalisation performance. The study mentioned in \cite{ma2022stateoftheart} involves a thorough examination of models for detecting microorganisms during analysis. A comparison of the latest approaches demonstrates that the You Only Look Once (YOLO) architecture is superior. Notably, YOLO's exceptional detection capabilities are attributed to its unique utilisation of the DarkNet19 backbone network and an original k-means clustering algorithm, which enables the accurate generation of anchor boxes.\par

The field of visual tracking has been significantly influenced by tracking-by-detection techniques. These approaches use feature learning and classifiers to differentiate cells from background components. Subsequently, they establish associations between cells across frames using probabilistic objective functions \cite{alkofani2006lineage}. These techniques are highly efficient in terms of computation and are reliable when dealing with low cell density. However, as the cell density increases, their effectiveness declines due to problems related to segmentation errors and distinguishing between cells. An alternative approach entails a five-step process for tracking that aims to overcome these limitations. This process includes cell detection, tracking, and motion filtering through the use of Interacting Multiple Models (IMM) filters \cite{li2008lineage} further compiled into a single, unified track. This approach specifically caters to the necessity of prolonged monitoring, which is essential for precisely locating and tracking cells in a populated environment.

**Methodology**

**3.1: Overview**

The methodology in this research paper takes a comprehensive strategy to systematically explore the usefulness of state-of-the-art machine learning approaches for the automated detection and tracking of intestinal stem cells (ISCs) in biomedical images. The framework is made up of two major components: (a) intestinal stem cell detection and (b) intestinal stem cell tracking. SOTA object identification techniques, YOLOv8 and YOLONAS (Neural Architecture Search), are developed in the former. YOLOv8 is highlighted in particular because the exceptional performance of YOLOv8 in accurately detecting objects at a microscopic level [20][23][25] makes it particularly noteworthy. Conversely, YOLONAS stands out for its specialised architecture tailored to detect smaller objects, such as ISCs in this specific case [26]. A notable feature of YOLONAS is the inclusion of an Automated Neural Architecture Construction (AutoNAC) system that optimises resource allocation during model training. The decision to prioritise these technologies is driven by a significant gap in current research regarding the utilisation of the YOLO framework and ByteTrack algorithm, particularly for identifying and monitoring ISCs [27].

In order to conduct an empirical analysis, various object detection models are trained with different configurations and architectural parameters. These models include YOLOV8-S, YOLOV8-M, YOLOV8-L, YOLONAS-S, YOLONAS-M, and YOLONAS-L. The training process involves using a 90\% dataset of ISCs images to ensure that the resulting model can effectively handle various ISCs structures. This approach aims to create a model that can generalise well in diverse scenarios. After completion of the training process, every model undergoes validation using a separate dataset that is 10\% ISCs images to assess its strength and applicability in various scenarios. The evaluation of the models involves considering performance metrics such as mAP50, mAP50-95 value and F1 score. The checkpoint with the best validation loss during training is saved and then incorporated into the ByteTrack algorithm. By utilising a supervision tool [28], this final algorithm not only tracks ISCs across consecutive images but also determines their directional movement and traces their real-time trajectory.

Section 3.2 explains the ensemble system structure, Section 3.3 describes the data available to undertake the research, and Section 3.4 describes the research approach to develop the system using the data.

**3.2: Ensemble System Structure Description**

Figure 1 illustrates the implementation of a model architecture that employs a multi-stage ensemble system strategy. This approach effectively combines diverse machine learning techniques to tackle the difficulties associated with detecting and tracking Intestinal Stem Cells (ISCs). In the initial stage, annotated and pre-processed images are used to train two object detection models: YOLOv8 and YOLONAS. Section \ref{ sec: methodologies\_the\_data\_data\_preprocessing } provides further details on this process. To optimise cell feature learning and ensure optimal accuracy, hyperparameter tuning is performed as discussed in Section \ref{ sec: methodologies\_research\_approch\_intestinal\_stem\_cell\_detection} of research findings. Using established metrics, a comparative evaluation of these models is conducted, and the best-performing models' checkpoints are preserved. These fine-tuned models serve as the foundation for detecting ISCs (Intestinal Stem Cells), which can identify cells in separate frames and provide the coordinates of their bounding boxes. After detection, the ByteTrack algorithm, which utilises a tracking-by-detection approach for Multi-Object Tracking (MOT), leverages the video sequence along with the trained object detectors to locate ISCs in each frame. The resultant detection boxes are then processed using Kalman filtering and the Hungarian algorithm to ensure accurate location predictions for new ISCs and simultaneously match predicted and actual tracks respectively. Each section of these established tracks forms the foundation for plotting ISCs trajectories, offering a valuable understanding of the route taken by these cells. The complete procedure results in the creation of a video that can be exported, displaying the real-time tracking of ISCs visually.

**3.3 The Data**

**3.3.1 Microscopy Dataset**

The Biosciences Department at Durham University generously provided the dataset used in this research. The dataset was developed using a confocal microscope, which is a specialised type of fluorescence microscope. Initially, the dataset was obtained in '.lif' format and it consists of 10 distinct series. Each series can be conceptualised as a video or sequence of images, with approximately 25 frames and dimensions of pixels 416 x 416. These frames display fluorescence images indicative of kinase activity in the gut and thus represent Intestinal Stem Cells (ISCs). Importantly, every series also includes Z-stacking, a digital image-processing method that combines multiple images captured at different focal distances to produce a composite image with improved depth of field. To improve ISC extraction in each frame, the selection of the Z value was manually selected for each series using the visualisation features provided by Fiji ImageJ software. Figure \ref{fig:dataset\_instances} presents a instance of randomly chosen frames from Series 1 and Series 7, giving an overview of the dataset. The directional arrows in Figure \ref{fig:dataset\_instances} (a) and Figure \ref{fig:dataset\_instances} (e) specifically indicate the movement direction of the Intestinal Stem Cells (ISCs) in the following frames within their respective series.\par

**3.3.2 Data Preprocessing**

The 'readlif' Python library was used to transform images from the '.lif' file into a Pillow object. Nevertheless, this specific dataset only permitted the extraction of 8-bit grayscale images using the library, and alternative libraries had restrictions on their ability to extract images. As a result, a Python script was created with the purpose of transforming grayscale images into green channel images. This conversion process aimed to retain the kinase activity present in the images and ensure their resemblance to the images in the dataset. Additionally, a separate Python script was generated using the OpenCV library to consolidate all extracted images into a sequence of frames for experimentation with the tracking algorithm.

Roboflow [30] was chosen for the purpose of image annotation because of its user-friendly interface. To expand the dataset, which initially consisted of only 142 images, image augmentation techniques were employed resulting in a total of 356 images, effectively increasing it by 150%. This augmentation is beneficial in preventing overfitting and aiding the model's understanding of the characteristics such as shape, size, and orientation specific to each individual ISC. These techniques included horizontal and vertical flipping, 90° clockwise and counter-clockwise rotation, upside-down flipping, -15° and +15° rotation, -36% and +36% saturation, -25% and +25% brightness, -25% and +25% exposure, up to 2px blur, and up to 10% pixel noise. The dataset was divided into three parts: 90% for training, 5% for validation remaining 5% for purposes.

Sample images from the training and validation datasets are illustrated in Figure /ref{fig:train/valid\_dataset\_instances}. In particular, Figure /ref{fig:train/valid\_dataset\_instances} (a) shows a training mosaic image containing annotated cells classified as class [0], with '0' representing the 'cell' class. On the other hand, Figure /ref{fig:train/valid\_dataset\_instances} (b) illustrates images taken from the validation dataset.

Research Approch

The study methodology used two state-of-the-art object detection models, namely YOLOv8 and YOLONAS, to address the task of automated identification of intestinal stem cells in microscopic pictures. In comparison to YOLOv8, the model presents a notable benefit in terms of computing efficiency and accuracy due to its compact dimensions. The architectural design incorporates a modified CSPDarknet53 backbone of YOLOv5, which consists of a total of 53 convolutional layers. Three separate iterations of YOLOv8, namely YOLOv8-S, YOLOv8-M, and YOLOv8-L, were selected and trained. Each version shown gradual improvements in both accuracy and parameter complexity. Significantly, the YOLOv8 models underwent mosaic augmentation, as seen in Figure \ref{fig:train\_valid\_dataset\_instances} (a). The approach entails the integration of four distinct pictures, which prompts the model to adjust to different item positions, levels of obstruction, and dissimilar pixel backgrounds. These characteristics have similar significance in the context of identifying intestinal stem cells. The training process consisted of 50 epochs, during which a small batch size was used. To address the issue of overfitting, a dropout rate of 40% was implemented. This was particularly important due to the restricted size of the training dataset. The learning rate used along with using the Adam Optimizer, were adaptively modified within the range of 0.0001 to 0.00001 based on the observed loss values. In contrast, the YOLONAS models were trained using non-mosaic pictures and were restricted to a maximum of 20 epochs, since further training beyond this point did not result in noteworthy improvements. The batch size and dropout rates were kept consistent with those used in the YOLOv8 models in order to mitigate the risk of overfitting. YOLONAS was assigned an initial learning rate of 0.0005, which was comparatively higher, to account for its unique architectural intricacies. The objective of the harmonised technique used in both the YOLOv8 and YOLONAS models was to develop a reliable and efficient system for automatically detecting intestinal stem cells in microscopic pictures.

In order to achieve the goal of precisely monitoring intestinal stem cells in microscopic pictures, this research utilises the ByteTrack algorithmic framework, as seen in Figure \ref{fig:framework\_architecture.png}. The distinguishing feature of ByteTrack lies in its ability to include bounding boxes with varying levels of confidence during the development of tracking trajectories. The object tracking system utilises pre-trained weights from the YOLOv8-M models as underlying structures for the purpose of detecting intestinal stem cells in consecutive frames. It then follows to extract necessary details, including the coordinates of the bounding box and the corresponding confidence values. After the detection process, the bounding boxes are categorised into high-level or low-level confidence categories, depending on a predetermined threshold value. In the first stage, bounding boxes with high confidence levels undergo processing using a Kalman filter to make predictions about the coordinates of the following bounding boxes. The predictions are next evaluated for their similarity score by comparing them with the ground truth, which is obtained using the Hungarian approach. Based on this score, the predictions are either preserved or dismissed, resulting in the formation of tracking routes with a high level of confidence. During Stage 2, boxes with low-confidence are subjected to a comparable analysis procedure in order to generate tracking paths. Tracks that fail to match the predetermined threshold for similarity score are methodically excluded from further evaluation. In the final stage, the coordinates of each bounding box that has been successfully monitored are used to plot the pathways that are traversed by the intestinal stem cells. The mapping process is implemented using the OpenCV library, providing a complete methodology that encompasses the detection and efficient tracking of intestinal stem cells in microscopic images.

To fulfil the objective of accurately tracking intestinal stem cells in microscopic images, this study employs the ByteTrack algorithmic architecture, as illustrated in Figure \ref{fig:framework\_architecture.png}. Unique to ByteTrack is its capability to consider both high and low-confidence bounding boxes in the development of tracking pathways. The object tracking system leverages saved weights from the YOLOv8-M models as foundational architectures to identify intestinal stem cells across frames and to subsequently extract pertinent data, such as bounding box coordinates and associated confidence values. Following detection, bounding boxes are classified into either high-level or low-level confidence categories based on a predefined threshold value. During Stage 1, high-confidence boxes are processed through a Kalman filter to predict the coordinates of subsequent bounding boxes. These predictions are then retained or discarded based on a high similarity score, determined using the Hungarian method, thereby forming high-confidence tracking pathways. In Stage 2, low-confidence boxes undergo a similar analytical process to create tracking pathways. Any tracks that do not meet the set threshold for similarity score are systematically eliminated from consideration. Ultimately, the coordinates of each successfully tracked bounding box are utilised to map the trajectories followed by the intestinal stem cells. This mapping is executed using the OpenCV library, thereby offering a comprehensive approach that not only detects but also proficiently tracks the movement of intestinal stem cells in microscopic imaging.

The F1‑score is usually applied to evaluate classification results as the harmonic mean of precision and recall.

Write the best quality academic scientific research paper-style Results and Evaluation section in a Paragraph.

Note: Adjust flow on your own and match all the context mentioned below  
By evaluating context of sentences please categorise it in Results and Evaluations

Below is the context of the flow I am expecting from the output.

Objective: Object detection model YOLOv8-S, YOLOv8-M, YOLOv8-L and YOLONAS-S, YOLONAS-M, YOLONAS-L was trained for the microscopic image of intestinal stem cells

This methodology is designed to offer a comprehensive framework for automated intestinal stem cell detection and tracking intestinal stem cell. This research shows new approach of using Deep Learning algorithm to use in study of intestinal stem cell. Novel YOLOv8-M model architecture is used to detect intestinal stem cells in images with mAP50 value of 90.73 and Tracking MOTP of 78.47% and MOTA of 76.63%.

Evaluation Metrics

Post-implementation, the performance of the detection and tracking components were quantitatively evaluated using metrics like precision, recall, F1-score for detection, and Multiple Object Tracking Accuracy (MOTA) for tracking.

Write the best quality academic scientific research paper-style Results and Evaluation section in a Paragraph.

Note: Adjust flow on your own and match all the context mentioned below

By evaluating context of sentences please categorise it in Results and Evaluations

Below is the context of the flow I am expecting from the output.

Evaluation for both Yolo models is required to choose outperforming models for required objective. In object Detection model like YOLO, mAP (mean Average precision) is a popular metric in measuring the accuracy of object detectors. Both models are validated on validation dataset during training to effectively chage training parameters during training and tested on unseen testing dataset to have equal testing grounds. This method will help in understanding true learning of features of object, in this case intestinal stem cells. Mean Average Precision (mAP) metric compares the ground-truth of the bounding box to the detected box and returns a score during evaluation. The higher the score, the more accurate the model is in its detections.

IoU (Intersection over Union) value indicates how better alignment between the predicted and actual regions. This measure helps to know if a region has an object or not. IoU is the ratio of the intersection area of two bounding box with area of Union of two bounding box. The IoU score 0.50 is good enough to conclude that the model detected the object correctly.

The evaluation of models are done on basis of mAP50 value which indicates the mean Average Precision calculated at IOU threshold 0.5 and mAP50-95 value indicates mean AP for IoU from 0.5 to 0.95 with a step size of 0.05. From Evaluation metrics of Table \ref{table:Per\_od} it can be observed that performance of YOLOv8 is superior than compared to YOLONAS value, mAP50, and mAP50-95 value of YOLOv8 model is 90.73 and 43.81 respectively, whereas other are lower than these value. Comparing YOLOv8 performance with other models YOLOv8-M model has outraging performance than other models which can be depicted from Table \ref{table:Per\_od} and Figure \ref{fig:graph\_eval\_od} (a) (b). F1 value is found to to be highest among other models of 0.81 from Table \ref{fig:Per\_od}.

Tracking algorithm is Evaluated on metrics of MOTP and MOTA value \ref{song2022evaluation}. MOTP (Multi-Object Tracking Precision) value is accuracy of localization of detection boxes similar to the mAP metrics having an ideal value of 1 or 100\%. MOTA (Multi-Object Tracking Accuracy) represent how many errors the tracker system has made in terms of Misses, False Positives, Mismatch errors. Tracking metrics 78.47% and 76.63 % of precision and accuracy from Table \ref{table:Per\_track} on testing dataset. Tracking results of Series 1 and Series 7 can be visualise in Figure \ref{ fig:results\_tracking}. Figure \ref{ fig:results\_tracking} (a), (b) represent instance of frame from Series 1 and Figure \ref{ fig:results\_tracking} (c), (d) represent instance of frame from Series 7. Closely observing Figure \ref{ fig:results\_tracking} (d), a interrupted tracks are illustrated in results is because of immediate transition of intestinal stem cells in video which enable to track cells effectively. From visual inspection it was found that immediate transition of cells in video fails to track them and effectively failing of tracking algorithm.

[1] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2965634/>

[2] <https://pubmed.ncbi.nlm.nih.gov/16574858/>

[3] <https://pubmed.ncbi.nlm.nih.gov/21376230/>

[4] <https://pubmed.ncbi.nlm.nih.gov/30753836/>

[5] <https://pubs.acs.org/doi/10.1021/acs.chemrev.8b00333>

[6] <https://pubmed.ncbi.nlm.nih.gov/23452846/>

[7] <https://www.nature.com/articles/nbt838>

[8] <https://www.nature.com/articles/nprot.2017.128>

[9] <https://pubmed.ncbi.nlm.nih.gov/30427308/>

[10] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2982878/>

[11] <https://pubmed.ncbi.nlm.nih.gov/35608229/>

[12] <https://pubmed.ncbi.nlm.nih.gov/27814364/>

[13] <https://www.nature.com/articles/nature14539>

**[14]** <https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>

[15] <https://arxiv.org/abs/1312.6229>

[16] <https://arxiv.org/abs/1411.4038>

[17] <https://arxiv.org/abs/1505.04597>

[18] <https://www.nature.com/articles/s41592-023-01879-y>

[19] <https://ieeexplore.ieee.org/abstract/document/9023664>

[20] <https://link.springer.com/article/10.1007/s10462-022-10209-1>

[21] <https://pubmed.ncbi.nlm.nih.gov/16434878/>

[22] <https://www.sciencedirect.com/science/article/pii/S1361841508000650>

[23] <https://arxiv.org/pdf/2304.00501.pdf>

(YoloV8)[25] <https://docs.ultralytics.com/models/yolov8/#usage>

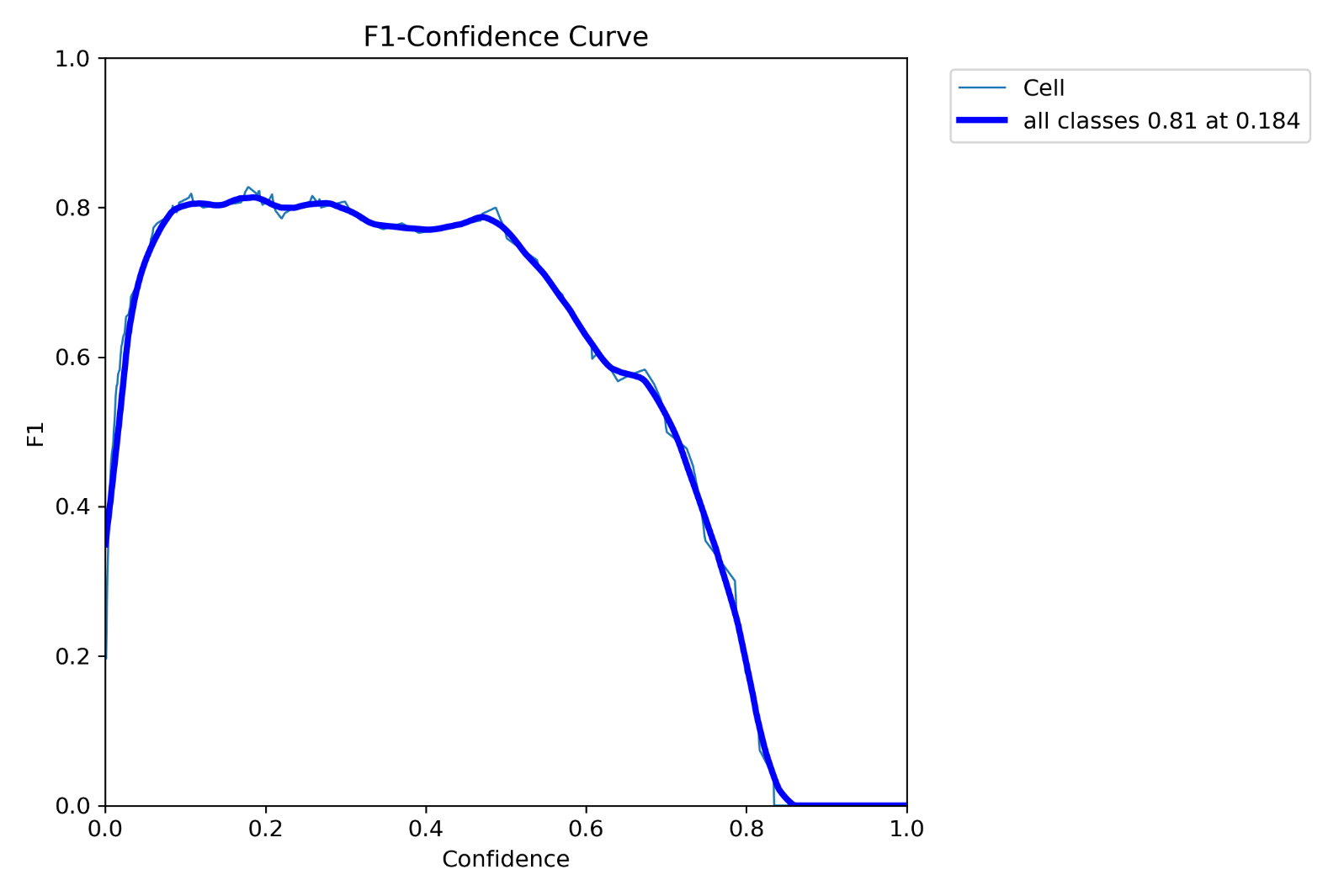
(YoloNAS)[26] <https://docs.ultralytics.com/models/yolo-nas/#citations-and-acknowledgements>

(ByteTrack) [27] <https://arxiv.org/abs/2110.06864>

(Supervision) [28] https://github.com/roboflow/supervision?ref=blog.roboflow.comsupervision%20citation

(Evaluation Metrics) [29] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8839404/

(Roboflow) [30] https://roboflow.com/research



**YOLONAS – S**

**YOLONAS – M**

**YOLONAS - L**

mAP50

Epochs