



Bioimage informatics

Digitally predicting protein localization and manipulating protein activity in fluorescence images using 4D reslicing GAN

Yang Jiao ¹, Lingkun Gu ², Yingtao Jiang ¹, Mo Weng^{2,*} and Mei Yang ^{1,*}

¹Department of Electrical and Computer Engineering, University of Nevada, Las Vegas, NV 89154, USA and ²School of Life Sciences, University of Nevada, Las Vegas, NV 89154, USA

*To whom correspondence should be addressed.

Associate Editor: Hanchuan Peng

Received on June 6, 2022; revised on September 28, 2022; editorial decision on October 17, 2022

Abstract

Motivation: While multi-channel fluorescence microscopy is a vital imaging method in biological studies, the number of channels that can be imaged simultaneously is limited by technical and hardware limitations such as emission spectra cross-talk. One solution is using deep neural networks to model the localization relationship between two proteins so that the localization of one protein can be digitally predicted. Furthermore, the input and predicted localization implicitly reflect the modeled relationship. Accordingly, observing the response of the prediction via manipulating input localization could provide an informative way to analyze the modeled relationships between the input and the predicted proteins.

Results: We propose a protein localization prediction (PLP) method using a cGAN named 4D Reslicing Generative Adversarial Network (4DR-GAN) to digitally generate additional channels. 4DR-GAN models the joint probability distribution of input and output proteins by simultaneously incorporating the protein localization signals in four dimensions including space and time. Because protein localization often correlates with protein activation state, based on accurate PLP, we further propose two novel tools: digital activation (DA) and digital inactivation (DI) to digitally activate and inactivate a protein, in order to observing the response of the predicted protein localization. Compared with genetic approaches, these tools allow precise spatial and temporal control. A comprehensive experiment on six pairs of proteins shows that 4DR-GAN achieves higher-quality PLP than Pix2Pix, and the DA and DI responses are consistent with the known protein functions. The proposed PLP method helps simultaneously visualize additional proteins, and the developed DA and DI tools provide guidance to study localization-based protein functions.

Availability and implementation: The open-source code is available at <https://github.com/YangJiaoUSA/4DR-GAN>.

Contact: mo.weng@unlv.edu or mei.yang@unlv.edu

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

Fluorescence microscopy, where samples are labeled with fluorescent probes (a.k.a. fluorophores), is one of the most versatile optical imaging methods. It allows visualization and quantification of various aspects of the target proteins, including their level, localization, behavior and interaction with other proteins. Usually, each protein of interest is labeled by one type of fluorophore, and its signals are collected as one channel. In laser-scanning confocal microscopy, labeled samples can be imaged in a 3D volume with three spatial axes. Application to live tissues or organisms generates time-lapse datasets with the time axis. Furthermore, if multiple proteins are labeled and imaged, additional channels are obtained, which result in information-rich 5D datasets.

Although multi-channel imaging is a powerful tool that is used to understand protein functions, in practice, the number of proteins that can be imaged simultaneously is limited. This is because the emission spectra of individual fluorophores are often too wide to be sufficiently separated. Additionally, the choice of fluorophores is limited by the quantum yield and photostability of fluorophores (Wall *et al.*, 2015), as well as the *in vivo* concentration of target proteins. When it comes to live imaging, the choice of fluorophores is particularly limited because signals from live samples are much weaker. These limitations contribute to the difficulties of simultaneously imaging more than two proteins in live samples. Beyond the choice of compatible fluorophores, the number of channels is also bound by the availability of laser lines and detectors of a microscope, the demand for acquisition speed, and the availability of

genetically labeled proteins. Without a proper tool, simultaneously observing and even studying multiple proteins has been quite a challenge.

One way to alleviate this challenge is to use machine learning methods to digitally predict the localization of unimaged proteins, using the localization information obtained from the imaged proteins. As a promising candidate model for this task, conditional generative adversarial networks (cGANs) are able to take an input image and generate the desired output image. A cGAN usually has a generator and a discriminator that are both convolutional neural networks. The generator uses network parameters to implicitly model the joint probability distribution of the inputs and the outputs so that it generates the desired output for any given new input. Theoretically, if enough samples and training time are offered, the modeled probability distribution can match the true distribution (Goodfellow et al., 2020). The discriminator model works as a classifier to discriminate the realness of the generated output. When given a new input, the generator tries to produce an output that fools the discriminator. In biological image processing, cGANs are popular in multiple topics including data augmentation (Bailo et al., 2019; Baniukiewicz et al., 2019; Dirvanskas et al., 2019; Osokin et al., 2017), domain translation (Han and Yin, 2017; Tang et al., 2020), resolution enhancement (Alam et al., 2021; Ishii et al., 2020; Wang et al., 2022; Zhou et al., 2020), virtual stain (Bayramoglu et al., 2017; Li et al., 2020; Liu et al., 2021; Rana et al., 2018; Rivenson et al., 2019; Vasiljević et al., 2021), stain normalization (Cong et al., 2021; Zanjani et al., 2018) and others (Isomura and Toyozumi, 2021; Kench and Cooper, 2021; Wang et al., 2021). Particularly, Pix2Pix (Isola et al., 2017) is a successful example of cGANs that show effectiveness on multiple tasks such as image colorization and style transfer. A recent work (Shigene et al., 2021) attempted to predict the localization of a protein using another protein from 2D fluorescence images with Pix2Pix.

However, the Pix2Pix work failed to obtain pixel-wise accurate results, likely because it only considered the 2D correlation between proteins. In living cells, the localizations of different proteins are often correlated in 3D space as well as in time. Due to the direct and indirect interaction, the localization of one protein complex may play a role in the localization of another complex of different protein compositions. Many proteins form large super-molecular complexes or structures, which occupy, move and interact with other complexes in 3D space that are captured in multiple z-slices of a 3D image stack. The complexes that are not in the same 2D plane may interact and provide important information for accurate localization prediction. In addition, complex formation, movement and interaction can be temporally regulated in the cell, and result in drastic localization changes over time. Thus, the localizations of interacting proteins often show temporal correlations as well.

To better incorporate 3D and time information in predicting the localization of proteins, we propose a protein localization prediction (PLP) method using a new cGAN named 4D Reslicing Generative Adversarial Network (4DR-GAN). 4DR-GAN models the joint probability distribution of imaged and unimaged proteins by incorporating the correlations between two protein localizations manifested in four dimensions, three in space and one in time. To our knowledge, this is the first work on applying cGANs to 4D information modeling. The generator of the 4DR-GAN is an end-to-end network that takes a 4D image as an input and incorporates spatial and temporal information via two encoding paths. Subsequently, the 5D feature maps are extracted from the two paths, and they are resliced to the same shape to be paired in space and time. The paired features are reconstructed to produce a 4D output image with identical size as the input. Altogether, this 4DR-GAN enables accurate prediction of protein localization that cannot be imaged together. Furthermore, with the new capability of accurate PLP of fluorescence images, it opens the door to digitally manipulating a protein's localization and activation. When manipulating the input protein, the response of the predicted protein reveals the protein relationships. In this regard, we further propose two novel tools, digital activation (DA) and digital inactivation (DI). DA allows to observe the predicted protein response when digitally increases protein localization or protein

activity, while DI serves the same purpose by digitally decreases protein localization or protein activity.

DA and DI present advantages when compared with genetic knockout and knockdown in terms of protein activity manipulation. Essential in testing the function of a gene, genetic knockout removes the gene from the genome, and gene knockdown stops or decreases the expression of the targeted genes. However, gene knockout and knockdown have drawbacks that mostly originate from their limited spatial and temporal control capabilities. Applying genetic knockdown or knockout to undesired tissues or stages often complicates the analysis of gene functions. Another drawback is that these genetic approaches are unable to manipulate gene functions at subcellular levels, which is an important aspect in understanding the differential protein functions at multiple subcellular compartments. In contrast, DA and DI can manipulate gene functions with precise spatial and temporal control and induce immediate effects, allowing gene function to be digitally removed or activated in any cells and subcellular regions at any time point. If the protein manipulation consistently leads to changes in prediction, the changes reflect the local or global relationship between the input and the predicted proteins, making DA and DI desirable tools for protein functional relationship study.

To evaluate the effectiveness of PLP along with DA and DI, we used 5D datasets from live imaging of *Drosophila* embryos that revealed the localization of two proteins in separate channels. These datasets offer rich temporal information since the subcellular localization of proteins change rapidly in a developing embryo. The high spatial and temporal resolutions of these datasets (pixel size: 0.1 μm ; frame rate: 10s) allow us to test prediction accuracy at subcellular levels. The proteins involved are well studied in their localizations and functions, and therefore, they offer a variety of evaluation criteria. We summarize our contributions in the following three aspects:

1. To visualize more proteins simultaneously in fluorescence microscopy, we propose a PLP method to predict the localization of unimaged protein from imaged proteins using 4DR-GANs, a new cGAN developed solely for this work. 4DR-GAN can simultaneously incorporate 4D information for the purpose of PLP.
2. Based on PLP, we developed two new tools to digitally manipulate protein localization and activation: DA and DI. These tools allow to precise spatial and temporal manipulation and induce an immediate response. A consistent response could provide clues to the functional relationship between the two proteins.
3. A comprehensive experiment on six pairs of PLP shows the effectiveness of 4DR-GAN and the success of PLP. Compared with the existing network, the protein localization and dynamic behavior in our prediction results are closer to the ground truth (GT). Through performing DA and DI on multiple groups of proteins, we obtained responses in the predicted protein localization that are consistent with the known protein functions.

2 Materials and methods

2.1 4D reslicing GAN

The fundamental role of 4DR-GAN is to incorporate spatial and temporal information simultaneously in the input 4D image and produce realistic 4D output. In PLP, 4DR-GAN takes one protein localization as input, and predicts another protein localization as output. 4DR-GAN consists of a 4D-reslicing generator (G) that predicts the protein localization in 4D images, and a 4D-consistency discriminator (D) that assesses the realness of the prediction in terms of localization, temporal consistency, and the input-target correlation. Figure 1 demonstrates the structure of 4DR-GAN, where an input 4D image is denoted as V_{α}^{xyz-t} , and a target 4D image is denoted as V_{β}^{xyz-t} . V_{α}^{xyz-t} is first resliced into XYZ-T view as V_{α}^{xyz-t} which sees the 4D image as XYZ-volumes with t frames, and XYT-Z view as V_{α}^{xyt-z} , which sees it as XYT-volumes with z frames.

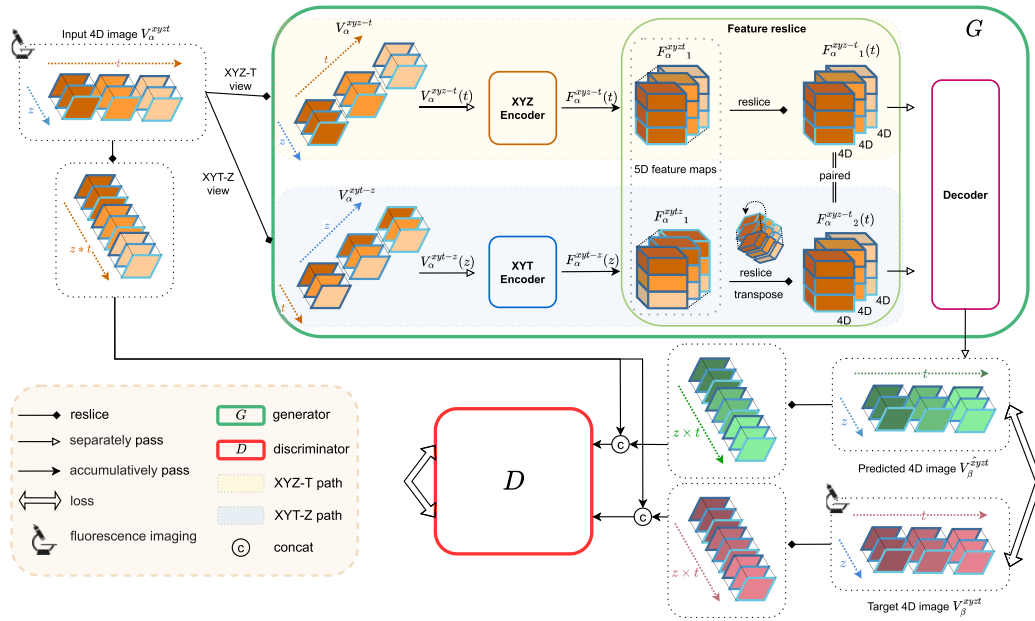


Fig. 1. 4D fluorescence image prediction by 4D Reslicing GAN (4DR-GAN). The overall flow of training the generator and the discriminator of 4DR-GAN. The input and the target 4D images constitute two channels of a 5D fluorescence image, which visualize the localization of two proteins in the X-, Y-, Z- and T-axes. G is a dual-path network that separately encodes the XYZ-axis and XYT-axis information of the input 4D image. D justifies the realism of the predicted image by taking the 4D images that are resliced into XY(ZxT) view. Various types of arrows are used to distinguish different operations, as shown in the legend. More details of network implementation are in [Supplementary Section 1.1](#) and [Figure S1](#). Training objectives and hyperparameters can be found in [Supplementary Section 1.2](#)

Correspondingly, the generator G has two paths. In the XYZ-T path of G , the XYZ-volume of each t frame $V_\alpha^{xyz-t}(t)$ is sent to XYZ Encoder to obtain the feature maps, denoted as $F_\alpha^{xyz-t}(t)$, while in the XYT-Z path, the XYT-volume of each z frame is sent to XYT Encoder to obtain the feature maps, denoted as $F_\alpha^{xyt-z}(z)$. $F_\alpha^{xyz-t}(t)$ and $F_\alpha^{xyt-z}(z)$ are 4D maps that incorporate both spatial and temporal information. All the feature maps, $F_\alpha^{xyz-t}(t)$ and $F_\alpha^{xyt-z}(z)$, are further assembled into 5D feature maps denoted as $F_\alpha^{xyz-t}_1$ and $F_\alpha^{xyt-z}_1$, respectively. Subsequently, taking into account that $F_\alpha^{xyz-t}_1$ and $F_\alpha^{xyt-z}_1$ represent different views of the image, we reslice $F_\alpha^{xyz-t}_1$ according to XYZ-T view to become $F_\alpha^{xyz-t}_2$, which spatially and temporally matches $F_\alpha^{xyt-z}_1$. This reslicing operation is detailed in the section of network implementation in [Supplementary Material](#).

To reconstruct a 4D output, the two 5D feature maps $F_\alpha^{xyz-t}_1$ and $F_\alpha^{xyt-z}_1$ in XYZ-T view will be independently decoded to obtain $V_\beta^{xyz-t}(t)$ in all t frames. Specifically, the 5D feature maps $F_\alpha^{xyz-t}_1$ and $F_\alpha^{xyt-z}_1$ are resliced into individual 4D feature maps $F_\alpha^{xyz-t}_2(t)$ and $F_\alpha^{xyt-z}_2(t)$ in t frames, and the corresponding pairs in time are sent to the Decoder to reconstruct the XYZ-volume $V_\beta^{xyz-t}(t)$. Upon the completion of the reconstruction, the prediction in all t frames V_β^{xyz-t} is obtained.

The discriminator D justifies the realism of the prediction by taking the 4D images that are resliced into XY(ZxT) view such as $V_\alpha^{xy(z \times t)}$, $V_\beta^{xy(z \times t)}$, and $V_\beta^{xy(z \times t)}$. D takes the localization of the input and the target proteins simultaneously, such as $D[(V_\alpha^{xy(z \times t)}, V_\beta^{xy(z \times t)})]$, $(V_\alpha^{xy(z \times t)}, V_\beta^{xy(z \times t)})]$. In this way, D justifies the localization and the temporal consistency of proteins, as well as the interaction between proteins.

2.2 Data acquisition

We applied this 4DR-GAN to the 5D datasets collected from live imaging of *Drosophila* early embryos involving three proteins: Myosin (Myo), E-Cadherin (E-Cad) and Ajuba (Jub). Embryos were dechlorinated in 4% sodium hypochlorite, washed in water, and mounted in glass-bottom Petri dishes by the natural affinity between the vitelline membrane and the glass. The dish chamber was then filled with water and covered by an oxygen-permeable membrane.

The imaging was performed with a Zeiss LSM 800 confocal microscope equipped with high sensitivity GaAsp detectors. The 488- and 561-nm lasers were used to excite GFP and mCherry, respectively. Images were acquired using a plan-Apochromat 63 \times /1.40 oil objective with the pinhole set at 1 Airy unit and the pixel size set at 0.124 μm . The z-stacks start from the embryo surface to 7 μm deep with 0.5 μm increments. The time interval between stacks is 10 s.

The original 5D fluorescence images have two channels and slight variations in shapes. The two channels are split as input and target, and then resized and cropped into training and testing samples. Accordingly, we performed four groups of PLP: from Myo to E-Cad, from E-Cad to Myo, from Jub to Myo, and from Jub to E-Cad.

For each group of proteins, 516 samples with a size of $256 \times 256 \times 16 \times 10 \times 2$ were used for training. Each sample had 256 pixels on the X- and Y-axis, 16 pixels on Z-axis and 10 time frames. Meanwhile, two channels were involved in each sample, where one channel was used as input, and another channel was as the target output (GT). These samples were cropped from 12 5D fluorescence images. For validation, 129 samples were cropped from another three images. Three images containing the Myo localization with a size of $256 \times 256 \times 16 \times 40 \times 1$ were used for testing.

3 Results

We evaluate the PLP from 4DR-GAN with three approaches and compare the results with Pix2Pix prediction. Section 3.1 demonstrates the similarity and difference between prediction and GT based on the characteristics of imaged biological structures. Section 3.2 quantifies the similarity between the prediction and GT images using Fréchet Inception Distance (FID). Section 3.3 quantitatively evaluates the protein distribution and behavior as well as their changes with time. Section 3.4 shows the application of PLP in DA and DI.

3.1 PLP accurately recapitulates protein localization at subcellular levels

Because the function and activation state of the protein determine the subcellular localization, we evaluated the similarity between the prediction and target GT, using key biological characteristics of the subcellular localization. Our datasets recorded the ventral cells of

fly embryos (Fig. 2a, light purple) during a period when these cells turned from a flat sheet into a tube-like structure. This tissue shape change is driven by the combined action of Myo and E-Cad (Fig. 2a). Myo is a molecular motor that generates the contractile physical force that changes cell shape while E-cad connects Myo filaments in individual cells into tissue-level network (Martin et al., 2009, 2010). During the imaging period, the amount of Myo proteins that are activated in ventral cells is increased, and the activation is restricted to the apical cortex of the cell beneath the cell membrane (Fig. 2a, red filaments). In confocal microscopy images, the activated Myo complexes are visualized as filamentous networks of high concentration, which appears in the top slides of the image stacks (Fig. 2a and c, top row input). The inactive pool of Myo appears to be uniform and at a low concentration since they diffuse freely in the cytoplasm of the cell. To apply the force generated by active Myo to change cell morphology, Myo networks in neighbor cells are connected through the interaction with E-Cad complexes. E-Cad complexes provide adhesions between neighboring cells. In the images, inactive E-Cad proteins uniformly diffuse on and label

cell membranes with low intensity. In contrast, activated E-Cad proteins that are engaged in cell adhesion are assembled into higher-order complexes, and appear in images as high-intensity clusters along the cell-cell boundaries (Fig. 2a and c, second row, target, 8×8 cells). By connecting to these E-Cad clusters, Myo filaments pull cell boundaries towards the center of the apical surface, therefore reducing cell apical surface areas (Fig. 2a). Meanwhile, in response to the force experienced by the E-Cad complex, Jub is recruited to the cell adhesion complex and detected as spots overlapping with a portion of E-Cad clusters along cell-cell boundaries (Rauskolb et al., 2014).

To evaluate the subcellular localization of the predicted proteins, we picked a z-slice close to the apical surface including major Myo and E-Cad signals and demonstrated the changes in consecutive t frames (Fig. 2b).

First, we compared the morphology of the protein localization. Figure 2c-e shows the PLP results, with four rows displaying the input protein localization, the GT of target protein localization and the prediction by Pix2Pix and our 4DR-GAN, respectively

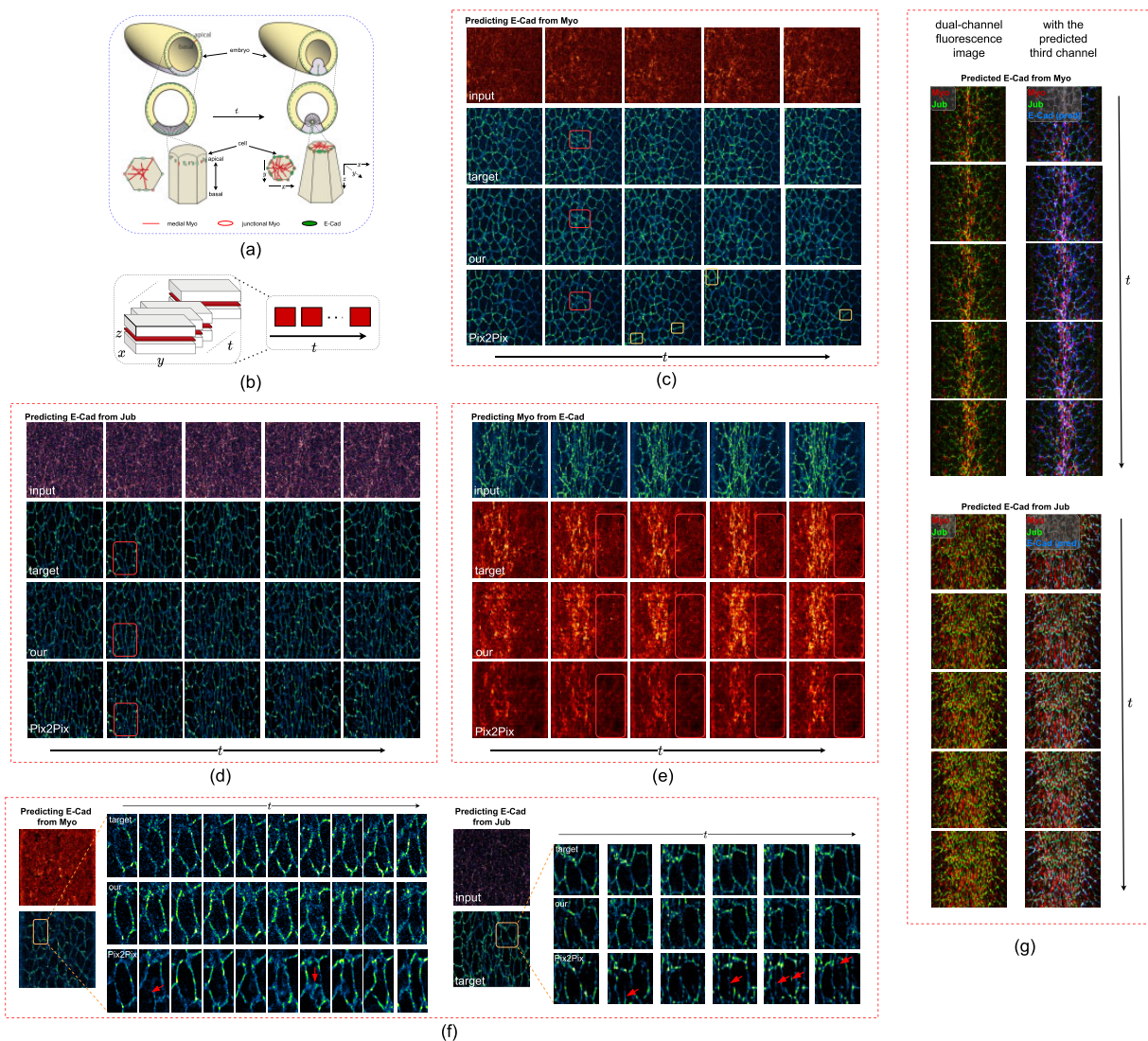


Fig. 2. The subcellular localization of the predicted proteins. (a) The known relationship between Myo and E-Cad. (b) The way of demonstrating 4D images with a z-slice and multiple time frames. (c) Localization prediction from Myo to E-Cad shown with the input Myo, the ground truth Ecad localization, our 4DR-GAN prediction results and the Pix2Pix prediction results. The red boxes show 4DR-GAN predicts better cell outlines and the yellow boxes show the prediction results are too smooth without showing E-Cad clusters. (d) Localization prediction from Jub to E-Cad. The red boxes show Pix2Pix predicts inaccurately resulting in extra cell boundaries. (e) Localization prediction from E-Cad to Myo. The red boxes show that Pix2Pix predicts inaccurately resulting in visible cell boundaries. (f) Temporal consistency of prediction results. Red arrows show inconsistent predictions in time. (g) Generating an additional channel that cannot be imaged together. Best view with zoom in (A color version of this figure appears in the online version of this article)

(additional cases in [Supplementary Figs. S5–S10](#)). As discussed above, E-Cad signals largely label cell boundaries. Consistent with the GT, 4DR-GAN produces correct outlines of individual cells, whereas extra or missing cells are often present in the Pix2Pix prediction ([Fig. 2c](#), red rectangle). 4DR-GAN is also better at recapitulating the clustering behavior of E-Cad proteins. Because inactive E-Cad uniformly labels cell membrane and active E-Cad form clusters, the cell outlines visualized by E-Cad are dotted lines like the target GT. By comparison, the Pix2Pix results tend to be smooth lines without clusters ([Fig. 2c](#), yellow rectangle). This shows that 4DR-GAN predicts the localization of activated E-Cad better than Pix2Pix most likely because 4DR-GAN utilizes the temporal information and allows capturing of more information from the predicted proteins. The advantage becomes more obvious when input images are of low signal-to-noise ratios, such as Jub channel co-imaged with E-Cad. As shown in [Figure 2d](#), the results produced by Pix2Pix show an extensive amount of extra cell boundaries that do not exist in the GT ([Fig. 2d](#), red rectangle). In contrast, our 4DR-GAN prediction is able to generate correct cell boundaries for most cells.

4DR-GAN also generates more faithful predictions of Myo from the E-Cad channel ([Fig. 2e](#)) and Jub channel ([Supplementary Fig. S8](#)). 4DR-GAN is able to recapitulate both the active Myo pool (high-intensity network) and the inactive pool (uniform at low intensity). Among the active Myo, the majority localizes in the center of the apical surface (medial Myo), while a minor pool localizes to some E-Cad complexes (junctional Myo). In Pix2Pix results, the predicted Myo shows lower intensities overall than GT and 4DR-GAN prediction. Interestingly, this inaccurate prediction affects medial Myo and inactive Myo more than junctional Myo, which results in a lower intensity ratio between medial Myo and junctional Myo in the Pix2Pix prediction. In addition, the localization of junctional Myo excessively resembles that of the input E-Cad rather than the GT Myo: cell outlines are clearly visible in Pix2Pix predicted Myo images even though cell outlines are barely visible in Myo images from GT and 4DR-GAN prediction ([Fig. 2e](#), red box). These analyses show that 4DR-GAN gives rise to more accurate prediction of protein subcellular localization.

Predicting Jub from E-Cad shows the expected cluster morphology along the cell boundary ([Supplementary Fig. S10](#)), but interestingly predicting Jub from Myo appears to be more challenging ([Supplementary Figs. S7 and S11](#)). It still predicts many aspects correctly. For example, Jub forms bright clusters similar to E-Cad, only in cells with high levels of active Myo, and in regions close to the apical cell surface. However, it could not precisely predict Jub localization to the cell boundary. This is surprising considering that E-Cad and Jub mostly localize together and predicting E-Cad from Myo is successful. One reason for the discrepancy may lie in the different reagents. Due to chromosome conflicts, the fluorescent Myo protein used in the Myo/Jub experiments is expressed from a different transgene than that used in the Myo/E-Cad experiments. The one used in Myo/Jub experiments appears to be expressed at a lower level. As a result, while junctional Myo is readily visible in Myo/E-Cad images, its signals are substantially weaker in the Myo/Jub images ([Supplementary Fig. S11](#)). This suggests that high-quality junctional Myo signals may be an important source of information for E-Cad prediction.

Secondly, we evaluated the temporal consistency of the predicted signals. Our 4DR-GAN is temporally more stable in terms of both pixel intensity and object morphology ([Fig. 2f](#)). In the GT and 4DR-GAN images, the pixel intensities of cell boundaries labeled by E-Cad are consistent between time frames. Whereas, in the predictions of Pix2Pix, the intensity of cell boundaries changes drastically, with some cell boundaries jumping from low to high intensity in a single time interval, only to drop in the next. Morphologically, it is observed that the shape of the same cell often changes sharply and cell boundaries can suddenly appear or disappear between time frames (red arrows in [Fig. 2f](#)). These predictions are incorrect as the shape and existence of cell boundaries do not change this drastically with our 10-s frame rate. 4DR-GAN reduces these problems and maintains the temporal consistency of cell morphology.

Lastly, to test the effectiveness of generating an additional channel that cannot be imaged together, we applied the trained 4DR-GAN to dual-channel datasets of Myo and Jub and generated E-Cad images as the third channel. Both Myo channel and Jub channel can successfully predict E-Cad channel, as shown in [Figure 2g](#). The predicted E-Cad gives rise to cell boundaries that not only are of appropriate sizes and shapes but also cover Jub signals, consistent with Jub localizing exclusively to a portion of the E-Cad complex. Lastly, consistent with the known spatial relationship of Myo and E-Cad localization, Myo appears mostly inside the cell boundaries labeled by predicted E-Cad.

3.2 4DR-GAN generates high-quality localization that has better FID scores than the compared baseline

To further evaluate the predictions generated by 4DR-GAN quantitatively, we employed the FID ([Heusel et al., 2017](#)). It is a widely used metric that reflects the human perception of similarity because it employs a deep CNN layer closer to output nodes that correspond to real-world objects. In contrast, the traditional pixel-level metrics, such as mean square error (MSE), structural similarity index and peak signal-to-noise ratio, are mismatched with human perceptual preference ([Zhang et al., 2018](#)). Two widely used pre-trained CNNs are employed in our evaluation: InceptionV3 ([Szegedy et al., 2016](#)) trained on ImageNet for image classification, and I3D ([Carreira and Zisserman, 2017](#); [Wang et al., 2018](#)) trained on Kinetics400 for video recognition. Because our prediction output is 4D, it is sliced on the Z-axis and T-axis into 2D images to fit InceptionV3 so that the similarity between prediction output and GT can be evaluated in the XY-plane. To fit I3D, the prediction output is sliced along Z-axis or T-axis, which results in the XYZ-volume and XYT-volume respectively. This allows the evaluation of volumetric and temporal consistency. We compared 4DR-GAN against Pix2Pix, and we further developed Pix2Pix from 2D to 3D to optimize its ability in PLP. More details of FID can be found in [Supplementary Section 1.3](#).

[Table 1](#) demonstrates the FID evaluation on the prediction of six groups of samples. To comprehensively evaluate 4DR-GAN as a dual-path network, an ablation study is conducted by separately evaluating the XYZ encoding path and the XYT encoding path while muting the fourth dimension. 4DR-GAN dual-path receives the best or the second best FID score in most evaluation cases than other networks. For the prediction quality in 2D, 4DR-GAN surpasses Pix2Pix as reflected on the FID with InceptionV3. This is consistent with the observation in Section 3.1 that the prediction of 4DR-GAN has more accurate cell outlines and protein clustering behavior. Since 4DR-GAN has a similar network depth and layer arrangement with Pix2Pix, the results show that predicting protein localization by incorporating 4D information helps improve the quality of prediction in 2D views. The FID with I3D(z) and I3D(t) further evaluates the volumetric consistency and temporal consistency, respectively, and in most cases, 4DR-GAN outperforms Pix2Pix. As expected, 4DR-GAN records a substantial improvement in temporal consistency because the temporal correlation is ignored in Pix2Pix. For example, in the case of predicting E-Cad from Myo, the score of volumetric consistency improves by 10.63%, from 0.762 to 0.681, and the score of temporal consistency improves even more, by 30.03%, from 1.612 to 1.130. Accordingly, the superior performance of 4DR-GAN on temporal and volumetric consistency supports the conclusion in Section 3.1 that the predictions of 4DR-GAN have stable pixel intensity and object morphology over time.

From the ablation study, overall, 4DR-GAN dual-path receives the best or the second best FID score in most evaluation cases: best score for 55.56% cases and the second best score for 33.33% of the case. When 4DR-GAN is the second best, its score is often very close to the best score and outperforms Pix2Pix. In addition, 4DR-GAN XYZ-path and XYT-path outperform Pix2Pix in most cases. When comparing 4DR-GAN XYZ-path and XYT-path with Pix2Pix, we observe that the prediction performance improvement is not constrained by the encoded dimensions in the generator. For example,

Table 1. Prediction quality evaluation by FID score

Protein (A→B)	Eval metrics: FID	Networks			
		Pix2Pix	4DR-GAN (XYZ path)	4DR-GAN (XYT path)	4DR-GAN (XYZ and XYT paths)
Myo → E-Cad	InceptionV3	2.243	1.320	0.932	<u>0.578</u>
	I3D (Z)	0.762	0.850	0.756	<u>0.681</u>
	I3D (T)	1.615	1.366	1.590	<u>1.130</u>
E-Cad → Myo	InceptionV3	2.872	0.631	0.524	<u>0.393</u>
	I3D (Z)	0.638	0.677	0.590	<u>0.557</u>
	I3D (T)	1.261	1.056	1.013	<u>0.882</u>
Jub → Myo	InceptionV3	1.861	0.610	0.377	<u>0.415</u>
	I3D (Z)	<u>0.920</u>	1.083	0.993	0.976
	I3D (T)	1.700	1.536	1.409	<u>1.379</u>
Myo → Jub	InceptionV3	3.083	1.867	1.903	<u>1.434</u>
	I3D (Z)	4.547	<u>1.397</u>	4.529	1.475
	I3D (T)	4.985	<u>1.775</u>	4.733	1.920
Jub → E-Cad	InceptionV3	4.253	1.404	<u>1.118</u>	1.157
	I3D (Z)	1.389	<u>1.205</u>	1.388	1.207
	I3D (T)	2.129	<u>1.547</u>	1.646	1.714
E-Cad → Jub	InceptionV3	5.689	1.717	2.090	<u>1.301</u>
	I3D (Z)	1.031	0.984	1.167	1.061
	I3D (T)	1.510	1.309	1.464	<u>1.264</u>

Note: Lower is better. The best results are in underline bold and the second best results are in bold. The orders of magnitude of InceptionV3, I3D (Z), and I3D (T) FID scores are 10^3 , 10^1 , 10^1 , respectively.

in the case of predicting Myo from E-Cad, we observe that 4DR-GAN XYZ-path performs well on both volumetric (XYZ) and temporal (XYT) consistency. The discriminator is the key to the improvement because it comprehensively justifies the localization and the temporal consistency of proteins, as well as the interaction between proteins. Although the 4DR-GAN XYZ-path focuses on encoding XYZ dimensions in the generator, the discriminator forces the generator to learn the T dimension to reduce the adversarial loss.

We observe that the score scales vary with networks used in FID. It is common for different pre-trained networks to result in scores that differ by a few orders of magnitude (Heusel et al., 2017; Wang et al., 2018). The biological images used in our study differ from the pre-trained datasets of Inception V3 and I3D, and result in different score scales. In addition, image defects such as inconsistent brightness and contrast across samples can affect the evaluation score scale as well.

The experimental result demonstrates that when two functionally related proteins are correlated in 4D, 4DR-GAN incorporates the information in all four dimensions and achieves high-quality prediction.

3.3 PLP predicts protein localization dynamics with high fidelity

Protein subcellular localization is dynamic during development and can change dramatically. We next evaluated the quality of the temporal dynamics of 4DR-GAN predicted protein localization. We used the Myo channel predicted from the E-Cad channel since Myo subcellular localization changes in all five dimensions during the live imaging periods. The details of evaluation implementation are in Supplementary Section 1.4 and Figures S2–S4.

First, we compared Myo intensity between GT and prediction in Z and T dimensions (Fig. 3a). Similar to GT, in the 4DR-GAN predicted channel, high-intensity Myo is only detected on the apical surface (the first several z slices) and becomes more and more intense during the imaging time frame. This is true for both medial and junctional Myo. Although the Pix2Pix prediction follows a similar pattern, the intensity of predicted Myo is lower. This is especially prominent for medial Myo, consistent with the 2D analysis that finds a lower medial: junctional Myo ratio (Fig. 2e). This becomes clearer when analyzing the intensity profiles along Z at a given time point or the intensity increase with time at a given z (Fig. 3b and c).

While GT and 4DR-GAN closely resemble each other, the profiles generated by P2P show lower intensities and somewhat deviated curves.

Secondly, we examined whether the predicted Myo localization is consistent with its biological function. Higher concentration (intensity) of filamentous Myo is correlated with higher contractile force (Xie and Martin, 2015). By connecting to E-Cad complex, the contractile force reduces the apical surface area. Therefore, at a given time point, cells with smaller apical surface area are more likely to have more active Myo. Figure 3d quantifies the average Myo intensities for cells of different sizes during the 10 consecutive time points when Myo is extensively activated. These histograms show that Myo is indeed at higher levels in cells of smaller size. The histogram profile of 4DR-GAN prediction is closer to that of GT than the Pix2Pix result. Similarly, a given cell usually has more active Myo when its area is reduced. Figure 3e shows the changing rate of Myo in cells of decreasing sizes at three time points. Within this short time frame (30 s), 4DR-GAN prediction and GT have over 30% of cells that show more than 10% increase in Myo. This parameter in Pix2Pix prediction is 22% which is 30–40% less than that of 4DR-GAN prediction and GT.

3.4 DA and DI predict correct consequences of protein loss-of-function and gain-of-function

The above analysis shows that, by integrating the information from all four dimensions, 4DR-GAN can accurately predict protein subcellular localization and concentration. Protein localization and concentration are not only the input and output of PLP, but are also closely related to protein activation states. For example, a higher concentration of Myo indicates more activated Myo proteins and correlates with higher physical tension generated by Myo. Therefore, we reason that it is possible to digitally control protein activities, by altering their localization and concentration in the images. This inspired us to develop effective DA and DI methods (refer to the DA and DI practice in Supplementary Section 1.6) to digitally manipulate protein activities and predict functional consequences. The predicted channel of 4DR-GAN-based PLP should respond to the input change in a way consistent with their functional relationship.

To test the effectiveness of these digital operations, we first used Myo as the input channel where Myo activities cause the change in

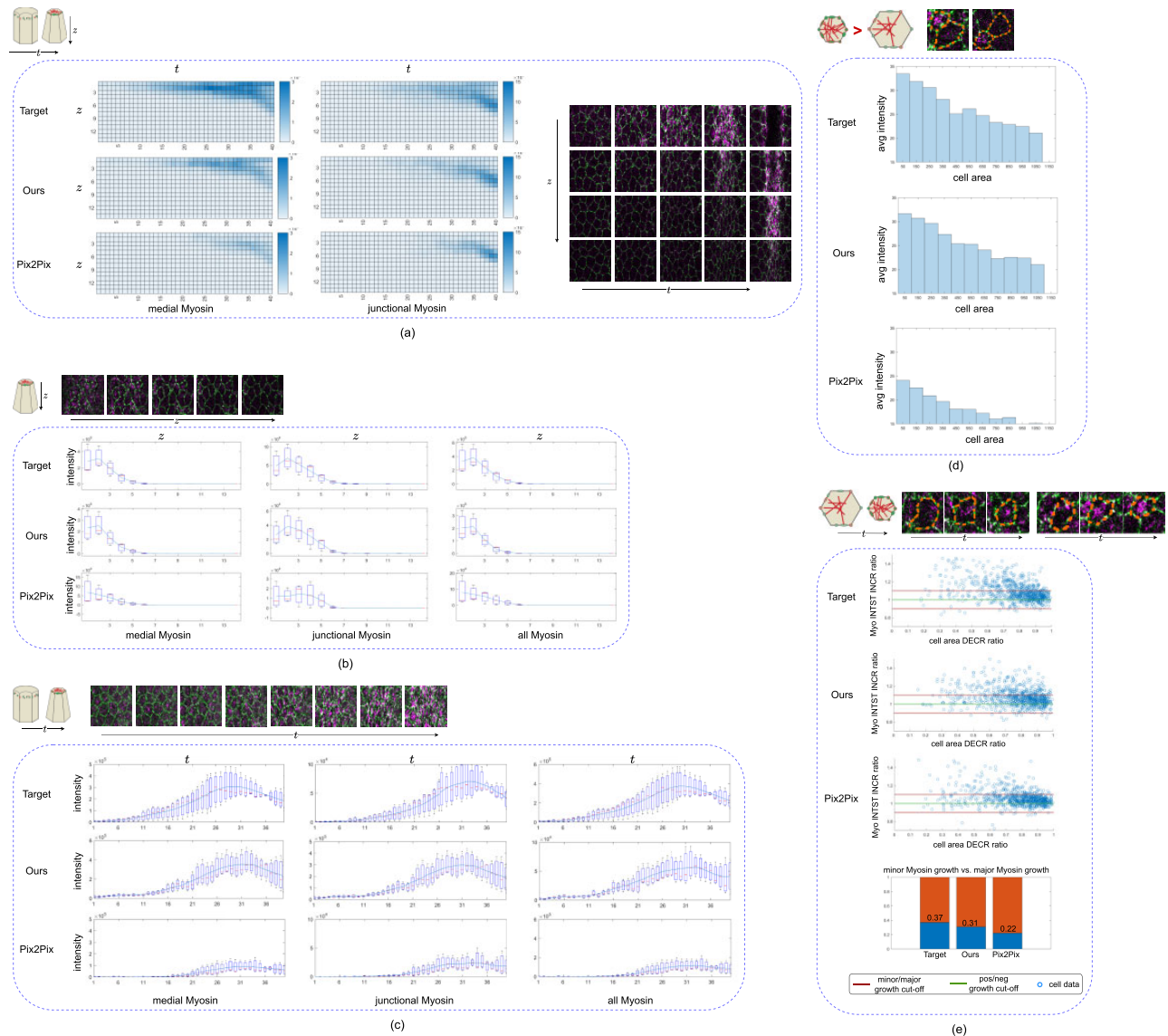


Fig. 3. The localization dynamic of the predicted Myo from E-Cad. (a) The predicted Myo intensity in Z and T dimensions compared with GT. Medial and junctional Myo signals are separately reviewed because they have differential relationships with the input E-Cad. (b) The predicted Myo intensity dynamics in Z dimension with a fixed time frame ($t = 30$) compared with GT. The example images show the GT signal in Z dimension. (c) The predicted Myo intensity dynamics in T dimension with a fixed z-slice ($z = 4$) compared with GT. The example images show the GT signal in T dimension. (d) The relationship between cell area and average Myo intensity in prediction results compared with GT. The example images demonstrate that the cells with smaller apical surface areas are more likely to have higher active Myo. (e) The relationship between cell area decreasing (DECR) ratio and average Myo intensity (INTST) increasing (INCR) ratio. The example images demonstrate that when the cell apical surface area decreases continuously, the cell is more likely to have increasing active Myo. Best view with zoom in

cell apical surface area. We performed DI of Myo in the circled region by erasing Myo signals (Fig. 4a, second row. More results in Supplementary Fig. S12). Since Myo produces contractile tension to reduce the cell apical surface, removing active Myo should lead to the relaxation of apical surface area, which appears in the image as bigger cells outlined by E-Cad. Indeed, an immediate response of the predicted E-Cad around the Myo knockout region is labeled: the cell outline marked by E-Cad in the new prediction becomes bigger (red) than those in the prediction before the DI (cyan). This is consistent with observations from biological loss-of-function experiments where breaking Myo filaments with high-power lasers leads to relaxation and expansion of the cell apical surface (Martin *et al.*, 2010). Figure 4b shows the effect of DA of Myo by increasing Myo intensity in the circle, which represents an increase in the contractile force and should lead to a reduction in the cell apical surface area. The cell outline marked by E-Cad in the new prediction becomes smaller (red) than that in the prediction before the DA (cyan).

Again, this is in line with the observations from biological experiments (Dawes-Hoang *et al.*, 2005), where forced Myo activation using genetic approaches induces cell apical surface area reduction.

In the above case, Myo is the driving force causing cell apical area to change. Next we ask, whether we can predict the localization pattern of the required Myo when we digitally manipulate the apical area. Specifically, we tested whether keeping cells from decreasing their apical areas would decrease the predicted Myo intensity and whether forcing cells to shrink would increase the predicted Myo intensity. Both operations generated the expected results around the digitally altered regions. In the result, compared with the prediction without manipulation, Myo is weaker when cell areas are kept from being reduced (Fig. 4c. More results in Supplementary Fig. S13), while Myo increases when one large cell is digitally split into two smaller ones that shrink in the apical area (Fig. 4d).

To test whether this approach can be applied to a variety of proteins, we digitally activated and inactivated Jub and observed how

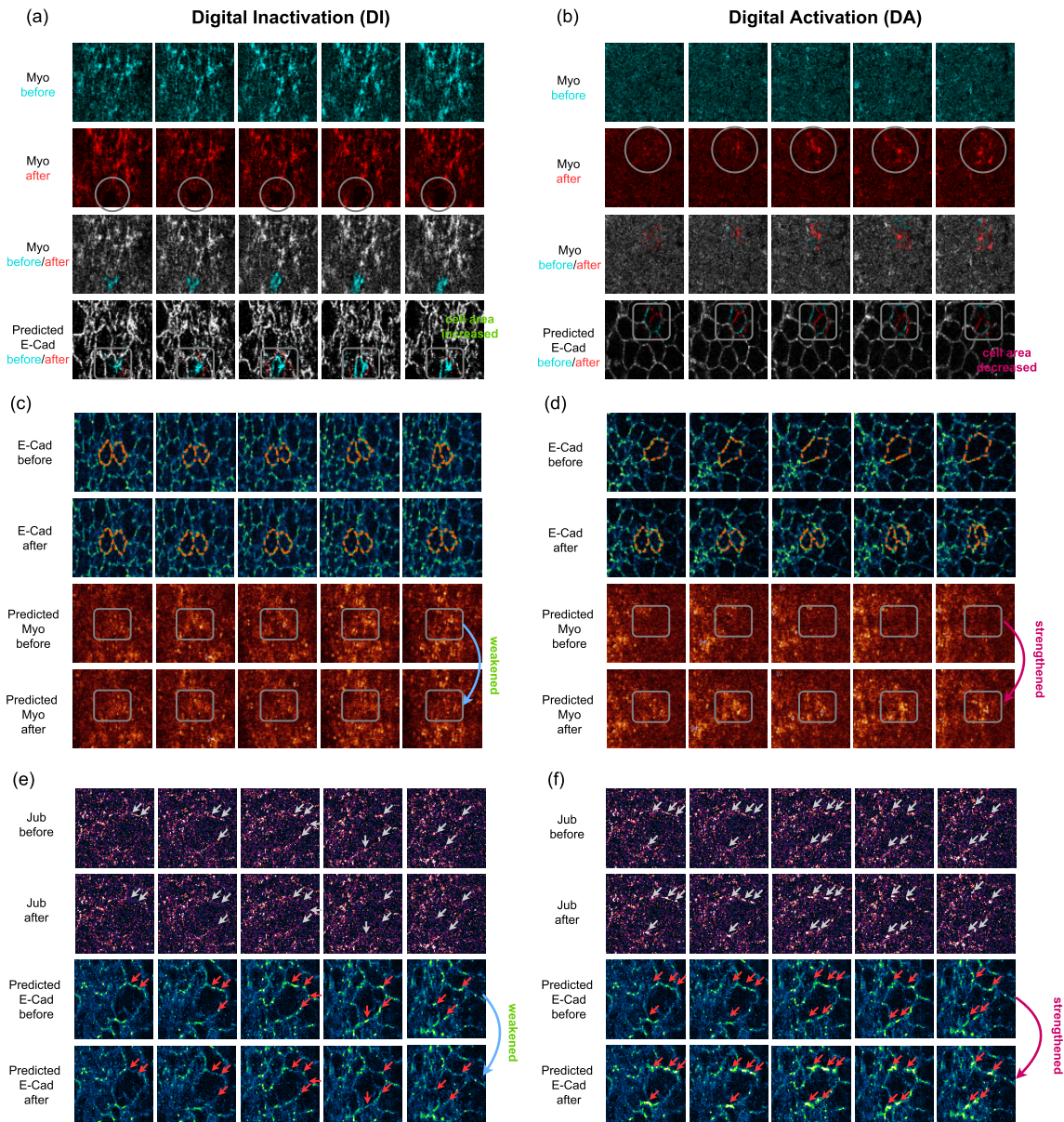


Fig. 4. Digital inactivation and digital activation results. (a) Observing E-Cad and cell response when digitally inactivating Myo subcellular localization. (b) Observing E-Cad and cell response when digitally activating Myo subcellular localization. (c) Observing Myo response when digitally inactivating E-Cad and stopping shrinking of apical areas. (d) Observing Myo response when digitally activating E-Cad and increasing shrinking of apical areas. (e) Observing E-Cad response when digitally inactivating Jub. (f) Observing E-Cad response when digitally activating Jub. The gray circles and arrows highlight the DI or DA manipulation, the orange dotted lines show the apical surface area manipulation, and the gray rectangles and red arrows highlight the responses. Best view with zoom in (A color version of this figure appears in the online version of this article)

E-Cad responds (Fig. 4e and f. More results in Supplementary Fig. S14). Jub diffuses in cytoplasm when inactive. It can be activated in response to Myo-generated tension and recruited to E-Cad clusters (Rauskolb et al., 2014). This recruitment of Jub into clusters is hypothesized to stabilize cell adhesion provided by E-Cad complexes (Razzell et al., 2018). Based on Jub localization properties, we digitally activated or inactivated Jub by strengthening or weakening Jub cluster intensity in the input images. It is observed that a weakened Jub cluster is translated into a weakened E-Cad cluster, and the two clusters overlap with each other. On the other hand, a Jub cluster of increased intensity is translated into a higher-intensity E-Cad cluster (Fig. 4e and f). This indicates that Jub and E-Cad not only colocalize in the images due to their molecular interaction but there is also a strong correlation between Jub and E-Cad clusters' intensity, consistent with Jub's role in stabilizing E-Cad-based cell adhesion (Rauskolb et al., 2014; Razzell et al., 2018). This also shows that

DA and DI are versatile approaches applicable to proteins with a network-like localization (Myo) and proteins with a cluster-like localization (E-Cad and Jub).

4 Discussion

Compared with Pix2Pix, PLP generated by 4DR-GAN is more accurate in subcellular localization, temporal consistency and dynamics. The experiment results demonstrate the importance of incorporating information from all spatial and temporal dimensions in the prediction of protein localization, which allows 4DR-GAN to capture more relationship features between two protein localizations. Noticeably, there are often different pools of the same protein that change with time and show differential localizations. When the pools of a protein play different roles in protein relationships, taking

advantage of four dimensions simultaneously is the key for accurate prediction. For example, predicting E-Cad from Myo localization requires the network to differentiate between junctional Myo and medial Myo. 4DR-GAN is able to do so and accurately predict the localization of E-Cad when junctional and medial Myo signals change dramatically with time. Similarly, the prediction from Jub to E-Cad requires 4DR-GAN to learn the relationship between Jub, activate E-Cad (high-intensity clusters) and inactivate E-Cad (low intensity uniform membrane). These experimental cases suggest that 4DR-GAN can learn complex spatial and temporal relationships.

The proposed PLP method will benefit a variety of fluorescence microscopies, especially live imaging where fluorophore choices are limited. For example, in our experiments, imaging Jub and E-Cad together is already challenging due to low *in vivo* protein concentration and fluorophore limitations, let alone imaging three proteins. With PLP, we successfully predicted E-Cad from the imaged Myo channel in Myo-Jub datasets, which leads to high-quality signals for all three proteins. PLP can also be instrumental in case of hardware limitations such as the availability of laser lines and detectors on a microscope by predicting other channels from imaged channels.

Based on 4DR-GAN-based PLP, DA and DI are two novel tools we propose for protein functional relationship study. A key feature of DA and DI is the capacity to precisely manipulate a protein localization in space and time. When doing so, DA and DI reflect the immediate effect on the output protein localization and therefore shed light on the protein relationships locally and globally. The experimental results not only demonstrate that DA and DI can predict the correct consequences of protein loss-of-function and gain-of-function, but also suggest that the 4DR-GAN-based PLP learns correct protein relationships. DA and DI require manipulation designs appropriate for protein functions. For unknown protein functions, multiple designs should be considered and analyzed.

Another key feature of DA and DI is that the outputs respond to input manipulation regardless of causality between proteins or the structure labeled by the protein. When the upstream protein is manipulated digitally, it mimics biological loss-of-function and gain-of-function experiments and the downstream protein responds in the predicted localization. However, the reverse experiment is different. When manipulation is applied to the downstream protein, the upstream protein does not respond in biological experiments, but it will respond in digital experiments. Specifically, when the downstream protein is digitally manipulated into a certain localization, the prediction provides a clue on the localization of the upstream protein localization that is required to drive the downstream protein into this certain localization. Therefore, DA and DI provide additional information in such cases. Overall, the realization of DA and DI provides a convenient and low-cost way to study protein functions and relationships and guides experimental designs in biological studies of unknown proteins.

Besides DA and DI, there are other possible applications of PLP. For example, in the cases of predicting the localizations of E-Cad from that of Myo and backward, we observe that the prediction of E-Cad from Myo has better perceptual quality than the prediction of Myo. The high prediction quality of E-Cad could imply that Myo is a major factor affecting E-Cad localization, while the information from E-Cad alone is insufficient and factors other than E-Cad contribute significantly to Myo localization. This observation also suggests potential future works of PLP. By adopting multiple informational sources, the performance of PLP for proteins that have multiple factors can be improved. In addition, analyzing the prediction quality may provide clues to the causality between proteins or cellular structures labeled by proteins.

Funding

This work was supported by the UNLV TTGRA and NIH Pathway to Independence Award [K99/R00 HD088764]. The publication fees for this article were partially supported by the UNLV University Libraries Open Article Fund.

Conflict of Interest: none declared.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

- Alam, M.S. *et al.* (2021) Super-resolution enhancement method based on generative adversarial network for integral imaging microscopy. *Sensors*, **21**, 2164.
- Bailo, O. *et al.* (2019) Red blood cell image generation for data augmentation using conditional generative adversarial networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA*, pp. 1039–1048.
- Baniukiewicz, P. *et al.* (2019) Generative adversarial networks for augmenting training data of microscopic cell images. *Front. Comput. Sci.*, **1**, 10.
- Bayramoglu, N. *et al.* (2017) Towards virtual H&E staining of hyperspectral lung histology images using conditional generative adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy*, pp. 64–71.
- Carreira, J. and Zisserman, A. (2017) Quo vadis, action recognition? A new model and the kinetics dataset. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6299–6308.
- Cong, C. *et al.* (2021) Texture enhanced generative adversarial network for stain normalisation in histopathology images. In: *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France*, IEEE, pp. 1949–1952.
- Dawes-Hoang, R.E. *et al.* (2005) Folded gastrulation, cell shape change and the control of myosin localization. *Development*, **132**, 4165–4178.
- Dirvanauskas, D. *et al.* (2019) HEMIGEN: human embryo image generator based on generative adversarial networks. *Sensors*, **19**, 3578.
- Goodfellow, I. *et al.* (2020) Generative adversarial networks. *Commun. ACM*, **63**, 139–144.
- Han, L. and Yin, Z. (2017) Transferring microscopy image modalities with conditional generative adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA*, pp. 99–107.
- Heusel, M. *et al.* (2017) GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: *Advances in Neural Information Processing Systems*, p. 30.
- Ishii, S. *et al.* (2020) Generative and discriminative model-based approaches to microscopic image restoration and segmentation. *Microscopy*, **69**, 79–91.
- Isola, P. *et al.* (2017) Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA*, pp. 1125–1134.
- Isomura, T. and Toyozumi, T. (2021) Dimensionality reduction to maximize prediction generalization capability. *Nat. Mach. Intell.*, **3**, 434–446.
- Kench, S. and Cooper, S.J. (2021) Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nat. Mach. Intell.*, **3**, 299–305.
- Li, D. *et al.* (2020) Deep learning for virtual histological staining of bright-field microscopic images of unlabeled carotid artery tissue. *Mol. Imaging Biol.*, **22**, 1301–1309.
- Liu, S. *et al.* (2021) Unpaired stain transfer using pathology-consistent constrained generative adversarial networks. *IEEE Trans. Med. Imaging*, **40**, 1977–1989.
- Martin, A.C. *et al.* (2009) Pulsed contractions of an actin–myosin network drive apical constriction. *Nature*, **457**, 495–499.
- Martin, A.C. *et al.* (2010) Integration of contractile forces during tissue invagination. *J. Cell Biol.*, **188**, 735–749.
- Osokin, A. *et al.* (2017) GANs for biological image synthesis. In: *Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy*, pp. 2233–2242.
- Rana, A. *et al.* (2018) Computational histological staining and destaining of prostate core biopsy RGB images with generative adversarial neural networks. In: *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA*, IEEE, pp. 828–834.
- Rauskolb, C. *et al.* (2014) Cytoskeletal tension inhibits Hippo signaling through an Ajuba–Warts complex. *Cell*, **158**, 143–156.
- Razzell, W. *et al.* (2018) The force-sensitive protein Ajuba regulates cell adhesion during epithelial morphogenesis. *J. Cell Biol.*, **217**, 3715–3730.
- Rivenson, Y. *et al.* (2019) PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl.*, **8**, 1–11.

- Shigene, K. et al. (2021) Translation of cellular protein localization using convolutional networks. *Front. Cell Dev. Biol.*, 9, 635231.
- Szegedy, C. et al. (2016) Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA*, pp. 2818–2826.
- Tang, Z. et al. (2020) 3D conditional adversarial learning for synthesizing microscopic neuron image using skeleton-to-neuron translation. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, IEEE*, pp. 1775–1779.
- Vasiljević, J. et al. (2021) Towards histopathological stain invariance by unsupervised domain augmentation using generative adversarial networks. *Neurocomputing*, 460, 277–291.
- Wall, K.P. et al. (2015) Fluorescence quantum yield measurements of fluorescent proteins: a laboratory experiment for a biochemistry or molecular biophysics laboratory course. *Biochem. Mol. Biol. Educ.*, 43, 52–59.
- Wang, T.-C. et al. (2018) Video-to-video synthesis. In: *Advances in Neural Information Processing Systems 31 (NeurIPS 2018), Montréal, Canada*.
- Wang, W. et al. (2022) Resolution enhancement in microscopic imaging based on generative adversarial network with unpaired data. *Opt. Commun.*, 503, 127454.
- Wang, Z. et al. (2021) Global voxel transformer networks for augmented microscopy. *Nat. Mach. Intell.*, 3, 161–171.
- Xie, S. and Martin, A.C. (2015) Intracellular signalling and intercellular coupling coordinate heterogeneous contractile events to facilitate tissue folding. *Nat. Commun.*, 6, 1–13.
- Zanjani, F.G. et al. (2018) Stain normalization of histopathology images using generative adversarial networks. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, IEEE*, pp. 573–577.
- Zhang, R. et al. (2018) The unreasonable effectiveness of deep features as a perceptual metric. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA*, pp. 586–595.
- Zhou, H. et al. (2020) 3D high resolution generative deep-learning network for fluorescence microscopy imaging. *Opt. Lett.*, 45, 1695–1698.