DELVING INTO THE EXPLAINABILITY OF PROTOTYPE-BASED CNN FOR BIOLOGICAL CELL ANALYSIS

Martin Blanchard[1], *Olivier Delézay*[2], *Christophe Ducottet*[1], *Damien Muselet*[1]

[1]Université Jean Monnet Saint-Etienne, CNRS, Institut d Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France,
[2]Université Jean Monnet, Laboratoire SAINBIOSE, INSERM UMR 1059

ABSTRACT

Deep learning for automated cell imaging analysis has become a tool of choice to process large amounts of data. But many of these methods lack explainability, slowing down their deployment for tasks such as diagnosis. We present a prototype-based framework to analyze structural changes which addresses the specific challenges of explainability in the context of cell imaging. Our method relies on classification between two distinct cell populations in a weakly supervised context where no label for individual cells is available. Our model extracts typical features from each population, representing intra-cellular structure, and provides an explanation on its classification decision by creating visualization of the local textures corresponding to the structures of interest. We show a real application where it effectively highlights a change in the organization of the actin content of the cells.

Index Terms— Visual Feature Learning, Biocellular Imaging, Explainable AI, Deep Neural Networks

1. INTRODUCTION

Early detection of disease is essential for successful treatment. In particular, this detection can be based on the analysis of biological cells to detect whether certain cells have a particular appearance (”degraded” or ”healthy”). Convolutional neural networks are a very powerful tool today for classifying biological cells from a labelled database. Unfortunately, this solution has two crippling drawbacks: the data must be precisely labelled and the network’s decision-making is rarely clearly explained to the domain expert, which is essential for human validation of the results.

In this paper, we propose a solution to extract and display structures that are typical of affected or healthy cells for a given treatment. For this purpose, we have grown cells in culture plates and exposed some of the wells in these plates to a product, such that we can extract cells from ”reference” wells (not exposed to the product) and from ”exposed” wells. The aim is then to develop a model that will enable us to analyse the cells in these wells and present the biologist with a set of ’prototypes’ representing healthy cells and a set of prototypes representing cells affected by the product. The first difficulty lies in the fact that we have no prior knowledge of the structural degradation of the cells induced by this product and we want this degradation to be provided by the network. This study is therefore part of the context of explainable AI. The second challenge is due to the fact that the cells are not labelled individually. Only the well is labelled (”reference” or ”exposed”) and we know that not all cells will react in the same way when a treatment is applied. Some cells will be effectively affected while others will not undergo any structural variation compared with their untreated state. This work is therefore part of a ”learning with noisy labels” context, since each cell in a well inherits the label of its ’exposed’ or ’reference’ well, even if the cell did not react to the treatment. The main idea of this work is illustrated in Fig. 1.

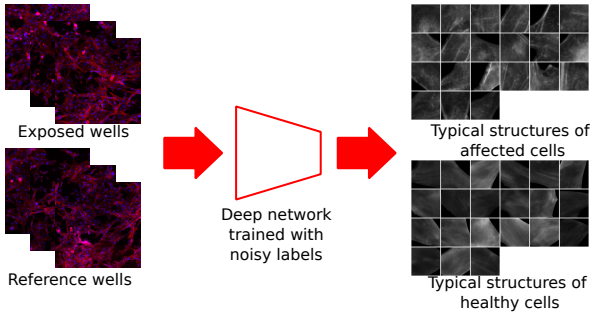


Fig. 1. The purpose of this work is to train a deep network with noisy labels (”Exposed”, ”Reference”) for each individual cell so that we are able to extract the typical structures of healthy and affected cells. These structures are illustrated by prototypes, exploited by the network during the classification task.

In this paper, we are taking advantage of the interesting results provided by prototype-based approaches for explainable deep networks [1]. Indeed, these solutions extract prototypical areas from images to explain their classification de-

This work has been funded by a public grant from the French National Research Agency (ANR) under the “France 2030” investment plan, which has the reference EUR MANUTECH SLEIGHT - ANR-17-EURE-0026

cision. ProtoPNet is a very representative example of these methods and has shown to provide excellent results for bird image classification [1]. Nevertheless, for the specific context of biological cells analysis, we show in this paper, that great care needs to be taken when choosing the design of several blocks. First, we create and provide masks to the loss function to avoid prototypes representing the background of the images. Second, we propose to adapt the receptive field in order to extract local texture features, that are meaningful for the biologists. Third, we freeze a part of the decision layer in order to control the prototype activation in the images. These points are discussed in detail in the next sections.

This paper is organized as follows. The next section is devoted to related works around deep explainable models and biological imaging. Then, we propose a mathematical formulation of our problem in section 3 and we present our method in section 4. Experimental settings and results are presented in section 5, and finally the paper is concluded with future work proposals in section 6.

2. RELATED WORK

As mentioned earlier, this work is related to both explainable deep networks and biomedical imaging.

2.1. Explainability methods

As explainability has always been a crucial challenge in deep learning, many methods try to explain decisions of a model. Traditional techniques include CAM [2] and Grad-CAM [3], which produce heatmaps highlighting zones of the input images that generate high activation of a given class. Saliency maps [4] are similar and allow to visualize which pixels contribute the most to the activation of a class. LIME [5] explains individual instances of predictions by training local interpretable surrogate model. But all of these methods are post-hoc explanations and do not make the internal reasoning process of the network interpretable. As explained in [6], trying to explain black box models rather than using inherently interpretable models can yield explanations that are unfaithful to the original model or not detailed enough.

Other techniques aim at understanding the convolutional filters learned by the models. Zeiler et al. [7] use a deconvolution approach to project the features back in the pixel space. Yosinski et al. [8] obtained similar visualization of the features by generating input images that maximize activations with regulations to ensure more interpretable features. Even if these techniques provide more insights into the model, they also are post-hoc methods and they might not be adapted to our cellular imaging problem where features are abstract and hard to analyze.

2.2. Applications of deep learning in biomedical imaging

Numerous attempts to use deep learning in the context of biomedical imaging have been made. Moen et al. [9] reviewed the different applications in the particular domain of cell imaging, and neural networks have proved to work very well for image classification. In another review focused on oncology, Jiang et al. [10] showed that deep learning can also be leveraged for diagnosis and evaluation of pathological features purposes. But they also underline the lack of interpretability of the techniques that are used. A third review [11] surveys the different types of interpretability methods that can be useful in the medical image analysis domain and their limitations. One of the reviewed articles [12] proposes an architecture combining a CNN classifier and a GAN discriminator to localize potential biomarkers in retinal or skin images with only image-level labels, in a context close to ours.

ProtoPNet [1] has also been used in the field of medical imaging. In [13], Mohammadjafari et al. used it to detect Alzheimer's disease in MRI scans from two different datasets while providing clear explanations as to why it classified inputs the way it did. Kim et al. [14] adapted the architecture for pathology diagnosis on chest radiographies by adding an occurence map to account for the position of the features in the input image

These successful works leveraged the classification capabilities of deep learning and ProtoPNet to allow for diagnosis of pathologies. Our goal is different as for us the classification is a pretext to understand the key visual features in the dataset and their biological meaning linked to the treatment used on the cells. As our work relies heavily on the use of the ProtoPNet model, we describe this approach in detail in section 4.1.

3. PROBLEM FORMULATION

We are conducting a biological experiment to investigate how a specific product influences the visual characteristics of a group of cells. We have two sets of cells: a reference group, denoted as $\mathcal{P}^{\mathrm{Ref}}$, which has not been exposed to the product, and an experimental group, denoted as $\mathcal{P}^{\mathrm{Exp}}$, which has been exposed to it. Initially, both sets of cells are cultured on plates and then treated with a labeling agent. Subsequently, they are examined under a microscope, with the images produced revealing the internal structure of each cell in response to the labeling agent. Each image contains numerous cells, each of which can be individually segmented. The image of a single cell is formally represented as I_i, and its corresponding segmentation mask is denoted as M_i, encoded as a binary image of the same spacial dimensions as I_i.

Our primary objective is to ascertain whether the cells exhibit visual changes due to exposure to the product. To achieve this, we assign a label, Y_i, to each cell i, where $Y_i = 0$ indicates no visual impact and $Y_i = 1$ indicates visual alter-

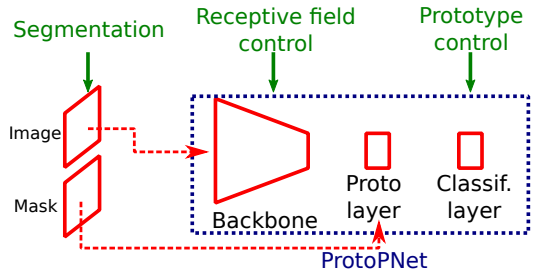


Fig. 2. Proposed framework based on ProtoPNet with three main adaptations: a segmentation mask as second input to prevent taking the background into account, a limited receptive field to ensure locality of the results and control of the classification layer weights to help achieve diversity in the prototypes.

ation. Thus, our task can be framed as a supervised classification problem, with the input data being $X_i = \{I_i, M_i\}$ for $i \in \mathcal{P}^{\text{Ref}} \cup \mathcal{P}^{\text{Exp}}$, and the goal being to predict the label Y_i.

However, from a biological perspective, we know that not all cells in $\mathcal{P}^{\text{Exp}}$ exhibit visual changes; some remain unaffected. Conversely, certain cells in $\mathcal{P}^{\text{Ref}}$ may display visual characteristics similar to affected cells. Furthermore, determining the true label Y_i is practically infeasible because we lack knowledge about the visual features of affected cells. Thus, our training is only weakly supervised, as we rely on imperfect labels derived from the population, which can be wrong at the cell level, denoted as $\widetilde{Y}_i = \delta(i \in \mathcal{P}^{\text{Exp}})$, while during testing, we aim to predict the true cell labels Y_i.

Given the unknown true labels, our objective is to analyze the decisions made by the classification system. This analysis serves two purposes: first, to evaluate the prediction quality, and second, to identify visual structures that are indicative of the visual appearance of affected cells. Specifically, we aim to learn prototypical structures that: i/ Characterize intracellular structures, ii/ Represent local textures and iii/ Serve as visual markers of the biological state of the cell.

4. OUR APPROACH

Our model is leveraging from classical prototype-based deep networks as illustrated in Fig. 2. On top of this architecture, we propose three important contributions. First, we apply a segmentation to the input image in order to extract individual cells and their associated mask. The mask is provided to the prototype layer. Second, we control the receptive field with an informed selection of the most appropriate backbone. And third, we freeze a part of the classification layer in order to control the prototype activation over the classes. These different steps are detailed below.

4.1. ProtoPNet

4.1.1. Principle

ProtoPNet is composed of three parts: a convolutional backbone f, a prototype layer g and fully connected layer h for classification. The backbone extracts relevant features from the input images I as a feature map Z. Then the prototype layer computes a similarity between this feature map and the prototypes with the following formula:

$$g_{p_j} = \max_{z \in \text{patches(Z)}} \log\left(\frac{d^2+1}{d^2+\epsilon}\right) \tag{1}$$

with g_{p_j} the prototype unit for a given prototype p_j, d the L^2-distance and z any patch of the feature map Z.

Finally, the fully connected layer multiplies the previous similarity outputs by its weight matrix w_h to obtain output logits for each class, which are then passed through a softmax function to be transformed into probabilities.

4.1.2. Training algorithm

The training of the architecture is decomposed in three stages. First, all layers but h are trained using the following loss:

$$\mathcal{L} = \frac{1}{n}\sum_{i=1}^{n} \text{CrsEnt}(h \circ g_{p_j} \circ f(X_i), \widetilde{Y}_i) + \lambda_1 \,\text{Clst} + \lambda_2 \,\text{Sep} \tag{2}$$

where CrsEnt designates cross entropy loss and n the training dataset size . Clst and Sep are defined as:

$$\begin{aligned} \text{Clst} &= \frac{1}{n}\sum_{i=1}^{n} \min_{p_j \in P_{\widetilde{Y}_i}} \min_{z \in \text{patches}(f(X_i))} d^2 \\ \text{Sep} &= -\frac{1}{n}\sum_{i=1}^{n} \min_{p_j \notin P_{\widetilde{Y}_i}} \min_{z \in \text{patches}(f(X_i))} d^2 \end{aligned} \tag{3}$$

The clustering term Clst pushes each training image to have one of its latent patches to be close to at least one prototype of the correct class. The separation term Sep pushes each training image to have all its patches to be far from any prototype that is not from the correct class.

The second step of the training algorithm is the projection of prototypes. During stage 1, the prototype layer was initialized with random points in the latent space that are then moved to minimize the loss.

Each prototype p_j is projected onto the nearest latent patch from a training image of the same class, in order for us to be able to understand what they visually mean. This projection process is defined as:

$$p_j \leftarrow \arg\min_{z \in \mathcal{Z}_j} d \tag{4}$$

with $\mathcal{Z}_j$ the set of all patches of same class as p_j. Since this operation moves the previously optimized prototypes, it lowers the classification accuracy. The last training stage consists of optimizing only the last layer to offset this accuracy drop.

4.1.3. *Prototype visualization*

During the projection stage of the training, each prototype is pushed to the closest latent patch of a training image. To find which patch in the original pixel space it corresponds to, Chen et al. forward the image from which the latent patch comes and upsample the activation map output from the prototype neuron g_{p_j} back to image size. The patch that represents the prototype is given by the smallest rectangle to include at least the 95th-percentile of all activation values.

4.2. Segmentation and masking

As stated in section 3, we want ensure prototypes represent intra-cellular structures. To this extent, we propose to use segmentation to work on individual cells. The segmentation mask can be used to restrict the model activation on non significant areas. The activation map output of the prototype layer is summed with the mask in order to artificially decrease the similarity between the background and the prototypes. Mathematically, the distance d between a patch $\tilde{z}$ and a prototype p_j becomes:

$$d = ||z - p_j||_2 + aM_z \tag{5}$$

where M_z is the value of the downsampled segmentation mask at the same coordinates as z and a is an arbitrary large number.

4.3. Receptive field control

Here we propose to limit the size of the receptive field to enforce the locality of the structures represented by the prototypes. A large receptive field will incorporate information from a large pixel-space patch into a single latent patch. Thus it becomes hard to understand what the similarity between the prototype and this latent patch actually means in the pixel space.

A smaller receptive field will help with locality as the corresponding pixel space patch is smaller. The deep features learnt by the convolutional backbone will more likely depict textures, rather than low abstraction-level structures as it usually is the case.

4.4. Prototype control

Given our data, what we expect to have in terms of prototype activations is illustrated in Fig. 3. For each prototype, we extract max similarity value from every image of the training set. Then we compute the mean value of these max similarities for separate populations $\mathcal{P}^{\text{Ref}}$ and $\mathcal{P}^{\text{Exp}}$ to obtain Fig. 3 (left). (right) is the difference between the two above distributions.

For cells from $\mathcal{P}^{\text{Ref}}$, the max similarity with Reference prototypes should be high and very low with Exposed prototypes. For $\mathcal{P}^{\text{Exp}}$ we expect high similarity with Exposed prototypes and a moderately high similarity with Reference prototypes as part of the cells can stay healthy.

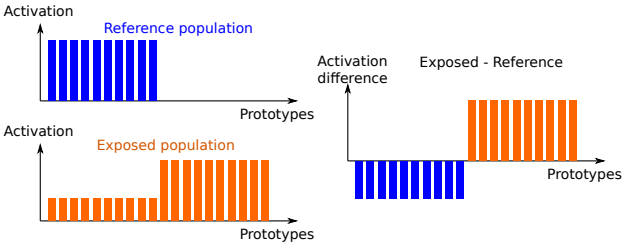


Fig. 3. Expected shapes of (left) mean max similarity over images of each population for each prototype and (right) difference of these mean similarities between Exposed (orange) and Reference (blue) populations.

We propose to freeze last layer weights with values 1 for weights linking one prototype to its correct class, and -0.5 for the opposite class, to help achieve these ideal similarities. Doing so helps the model to create prototypes that are primarily significant of their class while not completely removing the possibility of Reference prototypes being activated on Exposed cells as positive weights are higher than negative ones.

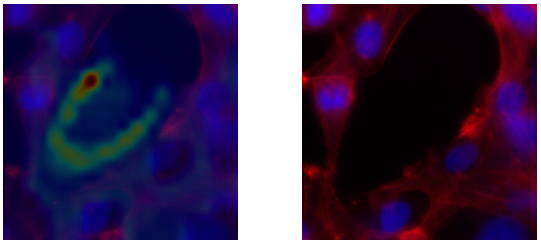
5. EXPERIMENTS

5.1. Dataset

Our dataset consists of over 3000 images of in-vitro endothelial cells from the blood-brain barrier. The blood-brain barrier (BBB) provides a tight seal between the bloodstream and the cerebral compartment, protecting brain from various potentially toxic molecules present in the systemic circulation [15]. We focused our analysis on the actin given its crucial role in the maintenance of the cell shape [16], as any modification of the filament structure could lead to disruption of the BBB.

Among the different factors that can influence the integrity of this barrier, TNF-α, an inflammatory cytokine, is known to lead alterations in the organization of junctional complexes [17]. Then, this cytokine could represent an interesting positive control compound for endothelial tissue disorganization.

Two populations were grown, one was treated with TNF-α at a concentration of 100 ng/mL (TNFA) and the other was not treated and serves as a reference to compare with (NT). Image acquisition was carried out with immunofluorescence on large portions of the wells, generating images of around 20.000×20.000 pixels. We used the red channel for the actin and the blue channel for nuclei. The nucleus channel is currently unused outside the segmentation process described in section 5.3. Channels were normalized separately and filtered for noise with a gaussian filter. The large images were then cut into smaller images of dimension 650×650 pixels. Data was split between train and test with 50% of images in each



(a) Upsampled activation map (b) Original image

(c) Upsampled activation map (d) Original image

Fig. 4. Example of the activation of a prototype created by the model on non-segmented input (b). The background patch activates the prototype (a). On a segmented input (d), the activation produced using the segmentation mask is located only inside the cell (c).

set. The training set was augmented offline two times with random rotation, flip, skew and shear applied.

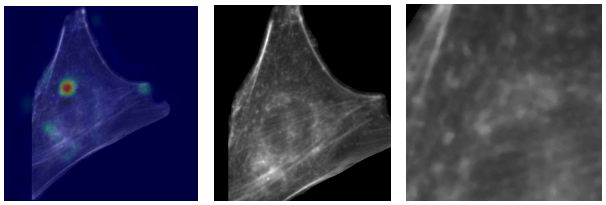
5.2. Experimental Setup

10 prototypes were created for each class. The factors λ_1 and λ_2, respectively multiplying the Clst, Sep of stage 1 loss were set to 0.8 and 0.08. The network was trained for 30 epochs, with the projection of prototypes happening every 10 epochs. These parameters will remain the same for all setups. With these default parameters and a DenseNet-121 [18] backbone we achieve accuracy of 0.7893 on our training set, without segmentation preprocessing.

5.3. Segmentation and masking influence

For the segmentation we used the Cellpose algorithm[19] pretrained on cytoskeleton images, using the actin channel as main channel and the nuclei channel as secondary. The result was not perfect, as we did not retrain their model on our data but it was sufficient to perform the rest of our analysis correctly. We will discuss later in section 6 of other ways to overcome this problem.

In Fig. 4 is a comparison of the activation maps produced by our model using non segmented input and activation produced on a segmented input, using the segmentation mask to condition the activations within the cell. We can see the non-segmented version of the model created a prototype that resembles the background which does not bear any biological information. This does not happen when using segmentation.



(a) (b) (c)

Fig. 5. Comparison of the prototype visualization for a given activation map (a) generated with using (b) the smallest box to enclose at least the 95th-percentile of activation, (c) a box of the size of the receptive field centered around the highest value of activation. Results obtained using VGG-11 backbone.

5.4. Receptive field influence

To validate our assumptions on the effect of the receptive field we tested two different convolutional backbones. We tried using DenseNet-121 [18] and VGG-11 [20] both pretrained on ImageNet [21]. DenseNet-121's receptive field is larger than our images while VGG-11 has a receptive field of 150×150 pixels.

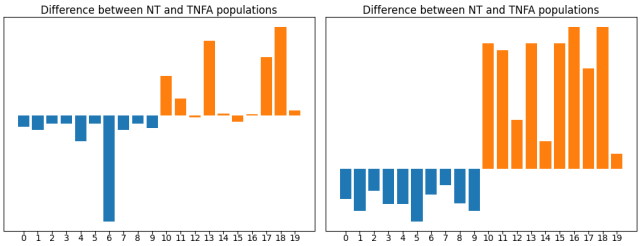
Fig. 5 shows that the initial method for visualization proposed within ProtoPNet is not adapted in our case as activation maps can have scattered peaks across the cell (a). Trying to enclose the 95th-percentile in a single box creates a box that almost contains the whole cell (b). Using a patch the size of the receptive field, centered around the max activation (c) gives a clearer vision of the texture that produced the maximum similarity with the prototype.

With VGG-11, it becomes obvious what the prototype encodes given the receptive field patch describes a single texture. As opposed, the corresponding receptive-field-sized patch we would have obtained with the DenseNet-121 backbone would have been the whole image and would not highlight any specific texture.

As with classification, the setup using DenseNet-121 achieved 0.8393 accuracy on test set, versus 0.7751 for the VGG-11 setup. This lower accuracy is not crippling as it still is reasonable and allows for extra interpretability as a tradeoff. Both setups use segmentation.

5.5. Prototype similarities

We described the ideal similarity distributions we aimed at in section 4.4. Here we show the mean max similarity difference between populations as in Fig. 3 (b). Fig. 6 shows the result obtained with optimizing the last layer as it is done in ProtoPNet (a) and with a frozen last layer (b). We can observe that the original method produces a very sparse distribution



(a) Non-frozen weights (b) Frozen weights

Fig. 6. Prototype similarities difference as in Fig. 3

with only a few prototypes that really bear resemblance with the cells. By freezing the last layer we have a denser distribution which means we capture typical structures within the cells with a greater variety. This property is especially useful in our case as we want to unravel the biological mechanisms behind the TNF-α action.

The control of the last layer brings the accuracy to 0.8975, much higher than previous setups and provides the best prototypes, as is illustrated in next section.

5.6. Biological Explanation

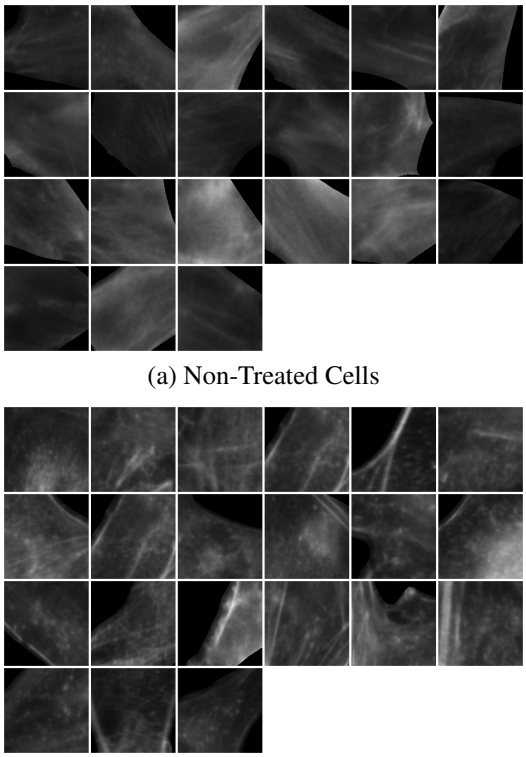
A single instance of a prototype is sometimes not enough to understand the underlying biological information as they depict textures. So we browsed the training set to retrieve the 20 images that activate the prototype the most and extracted the patches corresponding to the receptive fields. We stitched together these patches to obtain a mosaic that reflects the structure detected by the model as shown in Fig. 7.

In this form, we can see two different textures emerging. The first one is a rather uniform texture, with low contrast actin filaments. On the opposite, the second one is far more contrasted and shows a granulous texture along with actin filaments. From a biological standpoint, these two textures describe a modification of actin organization which could be linked to functional impairment of the endothelial cells and a disruption of the BBB.

6. CONCLUSION AND FUTURE WORK

We created a framework, based on ProtoPNet, that ensures explainability of its process. It is adapted to a cellular imaging problem as it extracts typical structures from the input cell populations by leveraging the power of deep learning for classification tasks, despite a weakly supervised context. The model provides visual elements that characterize intracellular structures with local textures, which can be analysed by an expert in the biological field to learn the biomarkers of a degradation associated with a given product.

Area for improvement include the segmentation process, which can be modified to only extract the background and



(a) Non-Treated Cells

(b) TNF-α treated cells

Fig. 7. Mosaics combining the visualization of one prototype and the 20 closest patches from the training set. (a) shows the result for a prototype of the non-treated class and (b) shows the result for a prototype of the TNF-α treated class.

work on patches of the cells as described in [22]. We also left aside the blue channel but meaningful information can be extracted from it [23]. Some hyperparameters have not been studied in depth, such as the number of prototypes per class, and could help improve the biological explainability.

7. REFERENCES

[1] C Chen, O Li, D Tao, A Barnett, C Rudin, and J K Su, "This looks like that: deep learning for interpretable image recognition," *Advances in neural information processing systems*, vol. 32, 2019.

[2] B Zhou, A Khosla, A Lapedriza, A Oliva, and A Torralba, "Learning deep features for discriminative localization," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, IEEE.

[3] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D, and Batra D, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *IEEE International Conference on Computer Vision, ICCV*

2017, Venice, Italy, October 22-29, 2017. 2017, pp. 618–626, IEEE Computer Society.

[4] K Simonyan, A Vedaldi, and A Zisserman, "Deep inside convolutional networks: visualising image classification models and saliency maps," 2014, pp. 1–8, ICLR.

[5] M T Ribeiro, S Singh, and C Guestrin, ""why should i trust you?": Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2016, KDD '16, p. 1135–1144, Association for Computing Machinery.

[6] C Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019.

[7] M D. Zeiler and R Fergus, *Visualizing and Understanding Convolutional Networks*, pp. 818–833, Springer International Publishing, 2014.

[8] J Yosinski, J Clune, A Nguyen, T Fuchs, and H Lipson, "Understanding neural networks through deep visualization," 2015.

[9] E Moen, D Bannon, T Kudo, W Graf, M Covert, and D Van Valen, "Deep learning for cellular image analysis," *Nature Methods*, vol. 16, no. 12, pp. 1233–1246, 2019.

[10] Y Jiang, M Yang, S Wang, X Li, and Y Sun, "Emerging role of deep learning-based artificial intelligence in tumor pathology," *Cancer Communications*, vol. 40, no. 4, pp. 154–166, 2020.

[11] Z Salahuddin, H C. Woodruff, A Chatterjee, and P Lambin, "Transparency of deep neural networks for medical image analysis: A review of interpretability methods," *Computers in Biology and Medicine*, vol. 140, pp. 105111, 2022.

[12] R Zhang, S Tan, R Wang, S Manivannan, Ji Chen, H Lin, and W Zheng, *Biomarker Localization by Combining CNN Classifier and Generative Adversarial Network*, pp. 209–217, Springer International Publishing, 2019.

[13] Mohammadjafari S, Cevik M, Thanabalasingam M, Basar A, and Alzheimer's Disease Neuroimaging Initiative, "Using protopnet for interpretable alzheimer's disease classification," in *Proceedings of the 34th Canadian Conference on Artificial Intelligence, Canadian AI 2021, online, May 2021*, Luiza Antonie and Pooya Moradian Zadeh, Eds. 2021, Canadian Artificial Intelligence Association.

[14] Eunji K, Siwon K, Minji S, and Sungroh Y, "Xprotonet: Diagnosis in chest radiography with global and local explanations," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021, pp. 15719–15728, Computer Vision Foundation / IEEE.

[15] M Blanchette and R Daneman, "Formation and maintenance of the bbb," *Mechanisms of development*, vol. 138, pp. 8–16, 2015.

[16] R Dominguez and K C. Holmes, "Actin structure and function," *Annual Review of Biophysics*, vol. 40, no. 1, pp. 169–186, 2011.

[17] R Versele, E Sevin, F Gosselet, L Fenart, and P Candela, "Tnf-α and il-1β modulate blood-brain barrier permeability and decrease amyloid-β peptide efflux in a human blood-brain barrier model," *International Journal of Molecular Sciences*, vol. 23, no. 18, pp. 10235, 2022.

[18] G Huang, Z Liu, L Van Der Maaten, and K Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, IEEE.

[19] C Stringer, T Wang, M Michaelos, and M Pachitariu, "Cellpose: a generalist algorithm for cellular segmentation," *Nature Methods*, vol. 18, no. 1, pp. 100–106, 2020.

[20] K Simonyan and A Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, pp. 1–14, Computational and Biological Learning Society.

[21] J Deng, W Dong, R Socher, L Li, Kai L, and Li F, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009, IEEE.

[22] G Campanella, M G. Hanna, L Geneslaw, A Miraflor, V Werneck Krauss Silva, K J. Busam, E Brogi, V E. Reuter, D S. Klimstra, and T J. Fuchs, "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nature Medicine*, vol. 25, no. 8, pp. 1301–1309, 2019.

[23] I Heckenbach, G V. Mkrtchyan, M B Ezra, D Bakula, J S Madsen, M H Nielsen, D Oró, B Osborne, A J Covarrubias, M. L Idda, M Gorospe, L Mortensen, E Verdin, R Westendorp, and M Scheibye-Knudsen, "Nuclear morphology is a deep learning biomarker of cellular senescence," *Nature Aging*, vol. 2, no. 8, pp. 742–755, 2022.