Supplementary Information

On the robustness of machine learning algorithms toward microfluidic distortions for cell classification via on-chip fluorescent microscopy†

Ali Ahmad [a,b] Federico Sala [c,d], Petra Paiè [d], Alessia Candeo [c], Sarah D'Annunzio [e], Alessio Zippo [e], Carole Frindel [b], Roberto Osellame [c,d], Francesca Bragheri [d], Andrea Bassi [c,d], and David Rousseau *[a]

1 LBP and LBP-TOP features extraction

Local binary patterns are among the most used texture descriptors in classification tasks[1]. For each central pixel position coordinate (x,y) of the image, the local binary pattern (LBP) indicates a sequential set of the binary comparison of its value with the P neighbors. The LBP assigns to each neighbor the value 0 or 1 if its value is smaller or greater than the pixel placed at the center, respectively. The resulting decimal value of the generated binary number replaces the central pixel value and it can be expressed as follows

$$LBP(x,y) = \sum_{n=0}^{P-1} 2^n b\ (i_n - i_{x,y})\ , \tag{1}$$

where $i_{x,y}$ is the gray value of the central pixel and i_n denotes the n^{th} neighboring one. Besides, the function $b(z)$ is defined as follows

$$b(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0\ . \end{cases} \tag{2}$$

The frequency of occurrences of each decimal code was then calculated over all the 2D microscopy images.

The local binary pattern of the 3D images was computed using the 3D extension of LBP, called LBP-TOP[2] originally applied for dynamic feature extraction in videos. This method only considers co-occurrence statistics in three planes XY, XZ, and YZ that intersect in the central pixel. LBP-TOP considers the feature distributions of each plane separately, then concatenates them together into a single histogram, making the feature vector shorter when the number of neighboring points increases, if compared to another LBP variant such as volumetric LBP (VLBP)[3]. Figure 1 shows the process of extracting features from the three orthogonal planes for a volumetric block and concatenating them into a histogram. The feature histogram for LBP-TOP can be defined as follows,

$$H_{i,j} = \sum_{x,y,z} I\{f_j(x,y,z) = i\},\ i = 0,...,n_j - 1\ , \tag{3}$$

where, n_j is the number of different codes produced by the LBP operator in the j^{th} plane ($j = 0$: XY, 1: XZ and 2: YZ), $f_i(x,y,z)$ expresses the LBP code of the central pixel (x,y,z) in the j^{th} plane, $x \in 0,....,X-1\}$, $y \in 0,...,Y-1\}$, $z \in 0,...,Z-1\}$ and $I\{A\}$ defined as follows,

$$I\{A\} = \begin{cases} 1, & \text{if A is true}\ , \\ 0, & \text{if A is false}\ . \end{cases} \tag{4}$$

Then, the calculated histogram was normalized to obtain a consistent description according to the following equation:

$$N_{i,j} = \frac{H_{i,j}}{\sum_{k=0}^{n-j-1} H_{kj}}\ . \tag{5}$$

[a] Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), UMR INRAE IRHS, Université d'Angers, 62 Avenue Notre Dame du Lac, 49000 Angers , France;
[b] Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé (CREATIS) , CNRS UMR 5220 – INSERM U1206, Université Lyon 1, Insa de Lyon, Lyon , France;
[c] Department of Physics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy;
[d] Istituto di Fotonica e Nanotecnologie, CNR, Piazza Leonardo da Vinci 32, 20133 Milano, Italy;
[e] CIBIO, University of Trento, 38123 Trento, Italy;
* Corresponding author: D. ROUSSEAU (e-mail: david.rousseau@univ-angers.fr).

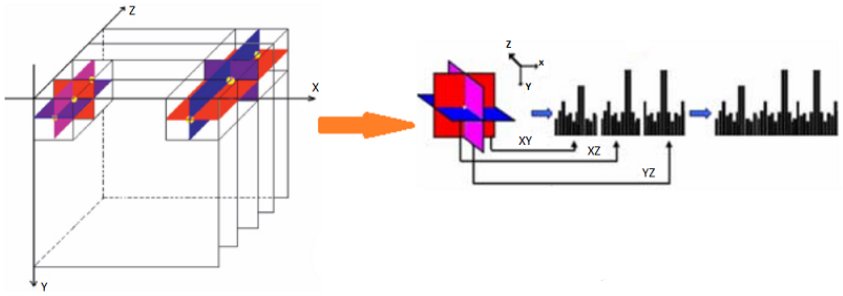

Fig. 1 Extraction of LBP-TOP features for each 3D block. Features are extracted from three orthogonal planes (XY, YZ, and XZ) and then concatenated together in one histogram.

2 Haralick features extracted from GLCM

From a Haralick co-occurrence matrix P of size n^2, we can calculate the following features that are most classically used:

1. Sum average:

$$\sum_{i=1}^{2n} iP_{x+y}(i) \tag{6}$$

2. Sum variance:

$$\sum_{i=1}^{2n} (i-\mu_{x+y})^2 P_{x+y}(i) \tag{7}$$

3. Contrast:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} (i-j)^2 P(i,j) \tag{8}$$

4. Correlation:

$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} ijP(i,j)-\mu_x\mu_y)}{\sigma_x\sigma_y} \tag{9}$$

5. Entropy:

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)log(P(i,j)) \tag{10}$$

6. Energy:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)^2 \tag{11}$$

7. Dissimilarity:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} |i-j|P(i,j) \tag{12}$$

8. Difference entropy:

$$-\sum_{i=0}^{n-1} P_{x-y}(i)log(P_{x-y}(i)) \tag{13}$$

9. Difference variance:

$$\sum_{i=0}^{n-1} (i-\mu x-y)^2 P_{x-y}(i) \tag{14}$$

10. Information measure of correlation 1 :

$$\frac{HXY - HXY1}{max(HX, HY)} \tag{15}$$

11. Information measure of correlation 2 :

$$\sqrt{1 - exp[-2(HXY2 - HXY)]} \tag{16}$$

12. Inverse difference :

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{P(i,j)}{1+|i-j|} \tag{17}$$

13. Sum entropy:

$$-\sum_{i=2}^{2n} P_{x+y}(i) log(P_{x+y}(i)) \tag{18}$$

14. Sum variance:

$$-\sum_{i=2}^{2n}(i-\mu_{x+y})^2 P_{x+y}(i) \tag{19}$$

In order to calculate the above characteristics, it is necessary to define the following statistics:

$$HX = -\sum_{i=1}^{n} P_x(i) log(P_x(i))$$

$$HX = -\sum_{i=1}^{n} P_y(i) log(P_y(i))$$

$$HXY = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j) log(P(i,j))$$

$$HXY1 = -\sum_{i=1}^{n}\sum_{j=1}^{n} P_x(i,j) log(P_x(i)P_y(j))$$

$$HXY2 = -\sum_{i=1}^{n}\sum_{j=1}^{n} P_x(i)P_y(j) log(P_x(i)P_y(j))$$

$$P_x(i) = \sum_{j=1}^{n} P(i,j)$$

$$P_y(j) = \sum_{i=1}^{n} P(i,j)$$

$$P_{x+y}(i) = \sum_{i+j=q} P(i,j), q = 2,3,2\dot{k}$$

$$P_{|x-y|}(i) = \sum_{i+j=q} P(i,j), q = 2,3,k\dot{-}1$$

$$\mu_x = \sum_{i=1}^{n}\sum_{j=1}^{n} iP(i,j)$$

$$\mu_y = \sum_{i=1}^{n}\sum_{j=1}^{n} jP(i,j)$$

3 Scattering transform (SCATNET)

A scattering transform defines a signal representation which is invariant to translations and potentially to other groups of transformations such as rotations or scaling. It is also stable to deformations and is thus well adapted to image and audio signal classification. A scattering transform is implemented with a convolutional network architecture, iterating over wavelet decompositions and complex

modulus.Figure 2 shows a schematic view of a scatter transform network working as a feature extractor and coupled to a classifier after dimension reduction. The scatter vectors Z_m at the output of the first three layers $m = 1,2,3$ for an input image f are defined by,

$$\begin{aligned} Z_1 f &= \{|f| * \phi\} \\ Z_2 f &= \{\cdots, |f * \psi_{j,\theta}| * \phi, \cdots\} \\ Z_3 f &= \{\cdots, ||f * \psi_{j,\theta}|\psi_{k,\varphi}| * \phi, \cdots\}, \end{aligned} \tag{20}$$

where the symbol $*$ denotes the spatial convolution, $|\cdot|$ stands for the L_1 norm, ϕ is an averaging operator, $\psi_{j,\theta}$ is a wavelet dilated by 2^j and rotated by θ. The range of scales $j = \{1,\cdots,J\}$ and the number of orientations $\theta = \{0,\pi/L,\cdots,\pi(L-1)/L\}$ are fixed by integers J and L. The number of layers is between $m = 1$ to $m = M$. A visualization of the Gabor filter bank used to extract SCATNET features and of the cell images at the output of the scattering transform array with 3 layers (m) for 4 scales (J) and 8 orientations (L) is shown in Figure 3.

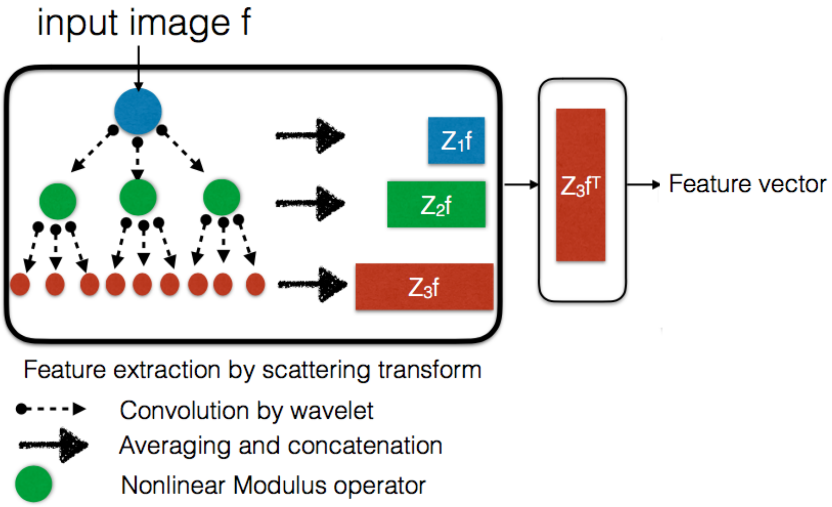

Fig. 2 Schematic layout of the cell images feature extraction based on the scattering transform with three layers. The feature vector at the output is the scatter vector $Z_m f$ of the last layer of Equation (20) after transposition.

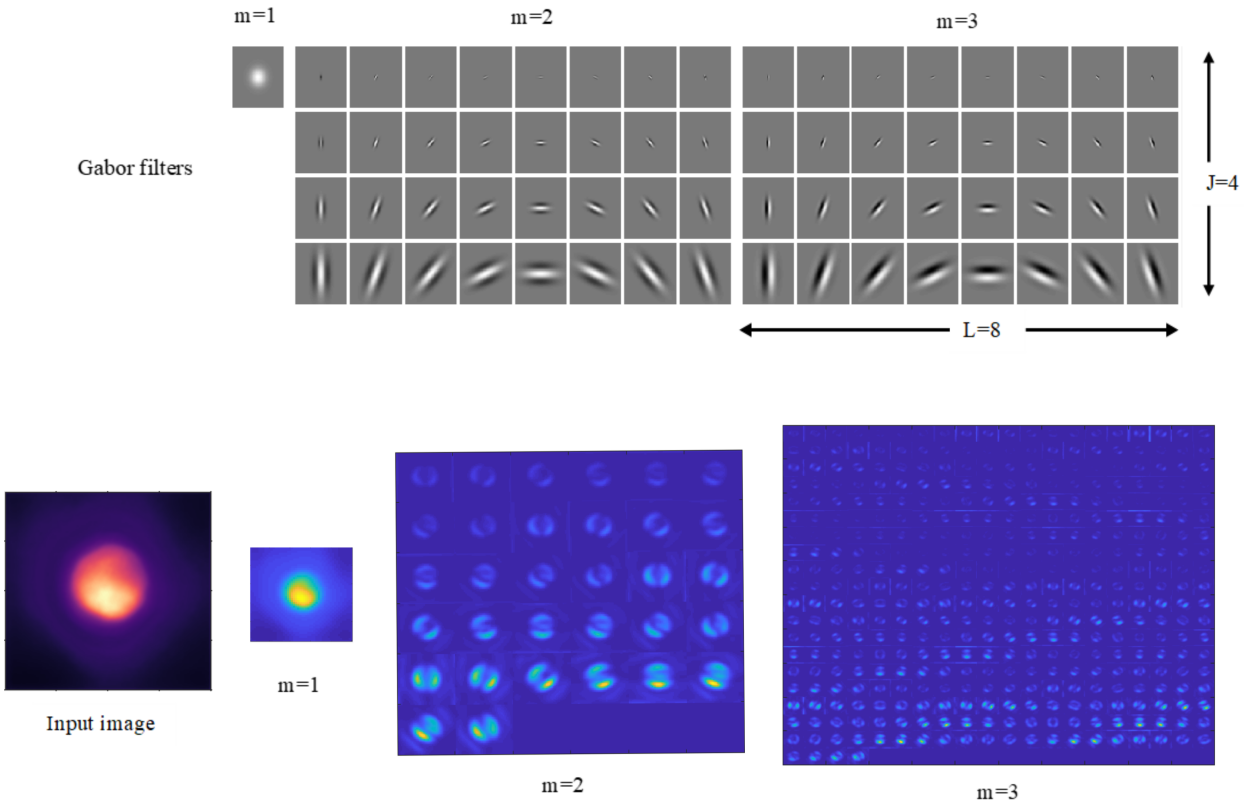

Fig. 3 Filter bank and images at the output of scattering transform network for each layer m of the scatter transform.

Notes and references

1 T. Ojala, M. Pietikäinen and T. Mäenpää, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, **24**, 971 – 987.

2 G. Zhao and M. Pietikainen, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2007, 915–928.

3 G. Zhao and M. Pietikäinen, *Dynamical Vision*, Springer, 2006, pp. 165–177.

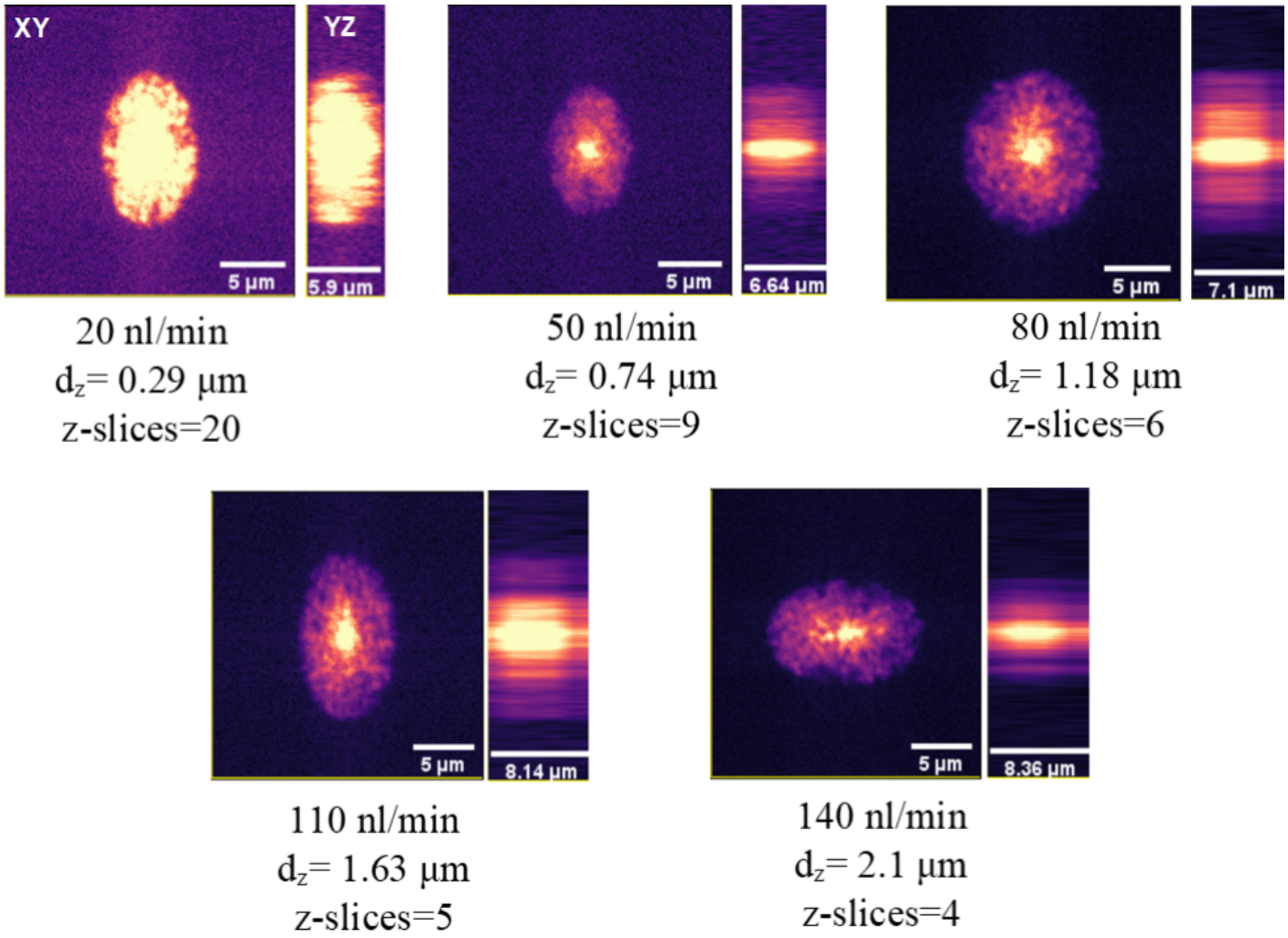

Fig. 4 Illustration of the XY and YZ planes located in the center of the simulated images as a function of the cell flow speed in the microchannels of the microfluidic system. The images shown are the XY and YZ planes of each simulated 3D image with the inter-slice distance (dz) and the number of slices acquired for each image for different cell speeds. The effect of microfluidics is characterized by the decrease in the number of z-slices and the increase of motion blur on the YZ planes when the cell speed increases.

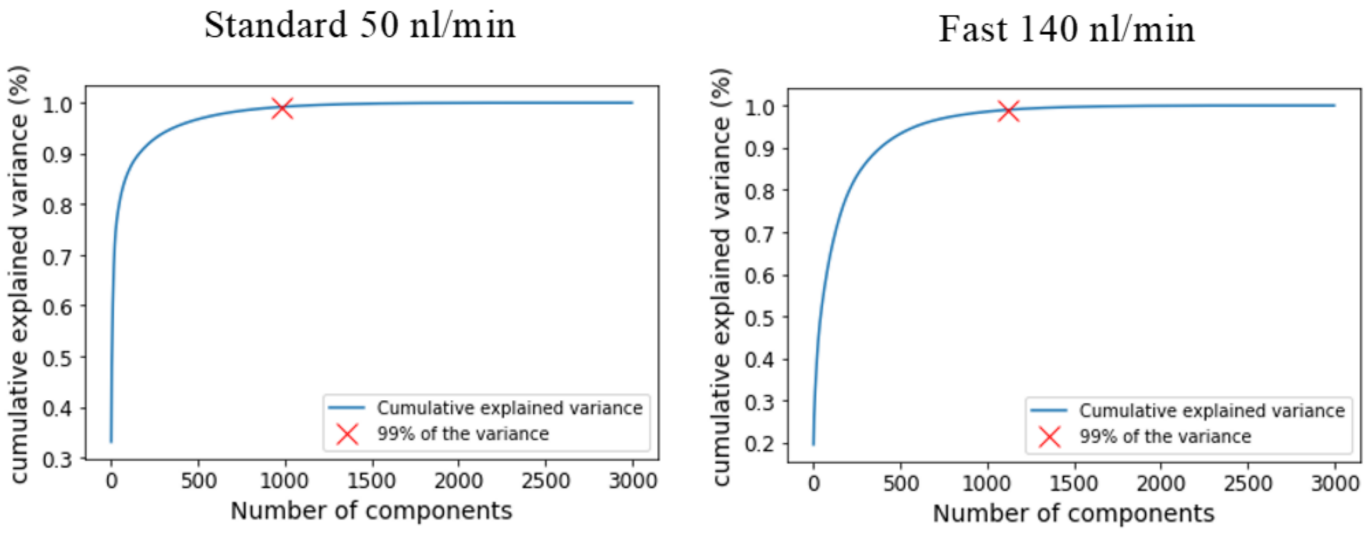

Fig. 5 Plots of the cumulative variance of the principal components function of the number of the components for 3D LBP features for both fast and standard cell flow speed cases. The red crosses corresponds to the number of components of 99% of variability (for standard speed: 963 components, for fast speed: 1081 components). The number of components used in principal component analysis of 3D LBP feature is set to 1000 for both cell velocities.

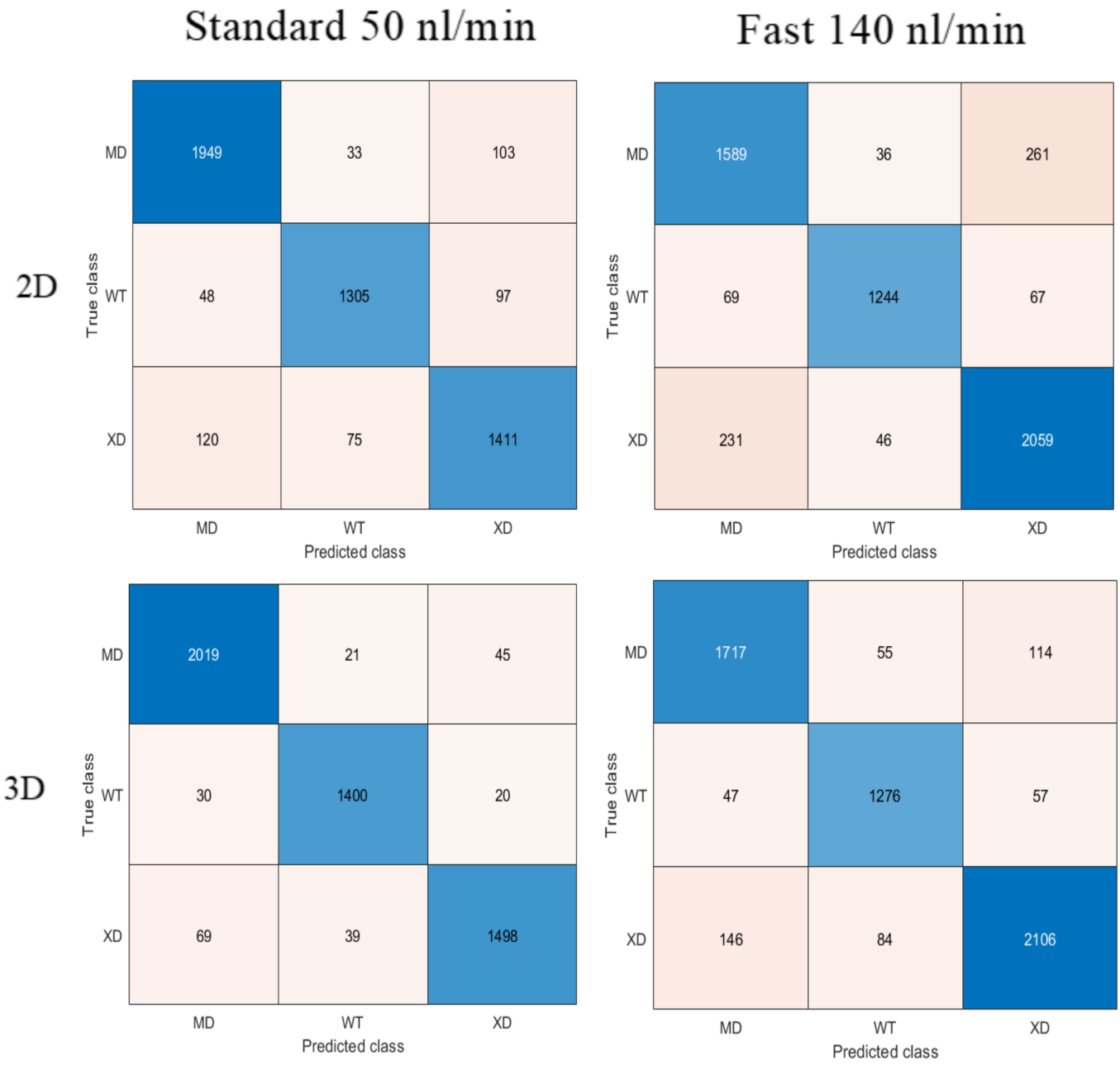

Fig. 6 Confusion matrices for the cells classification for standard and fast cell velocities based on 2D/3D LBP features computed from $10-$folds cross validation method.

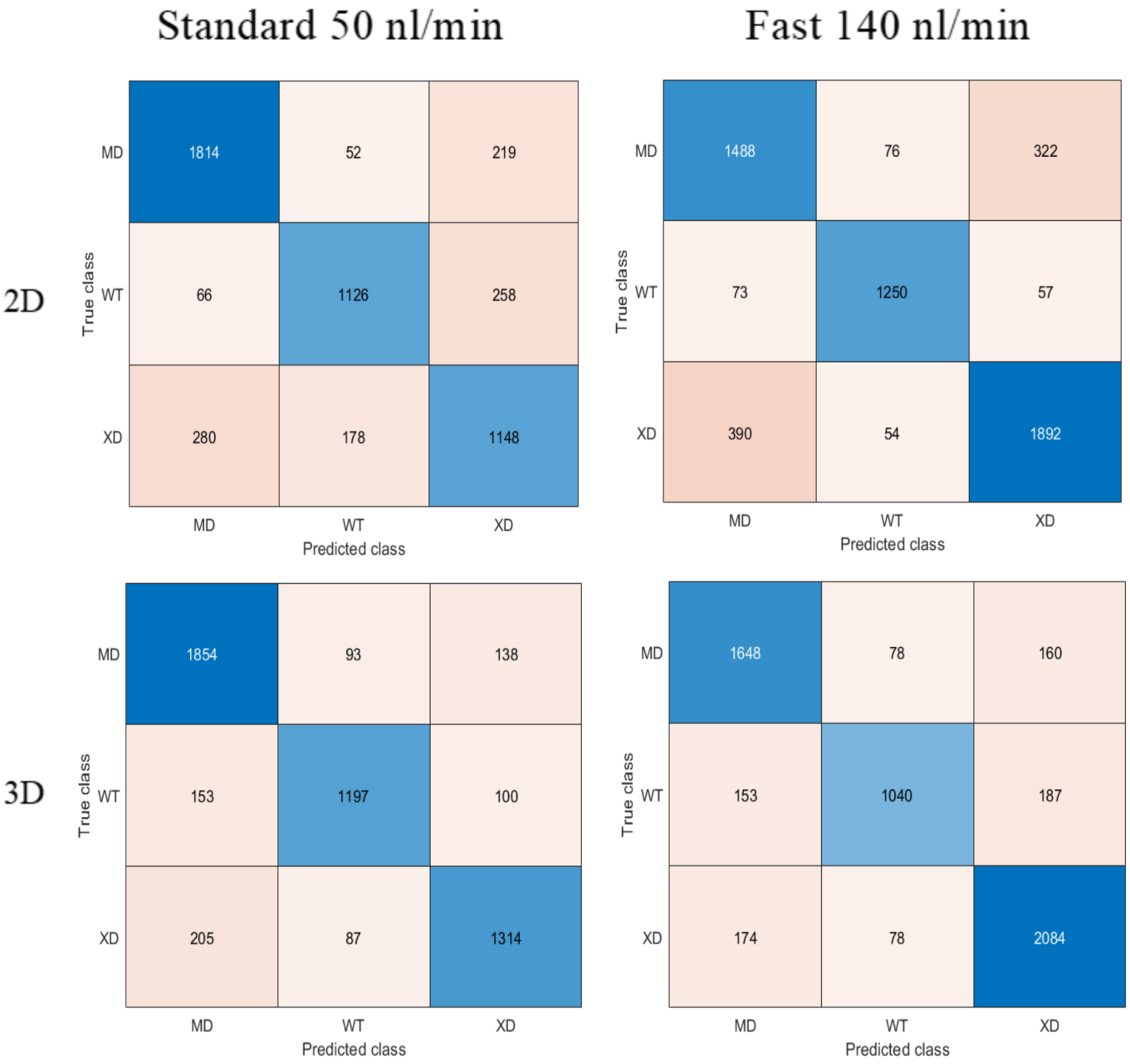

Fig. 7 Confusion matrices for the cells classification for standard and fast cell velocities based on 2D/3D GLCM features computed from 10−folds cross validation method.

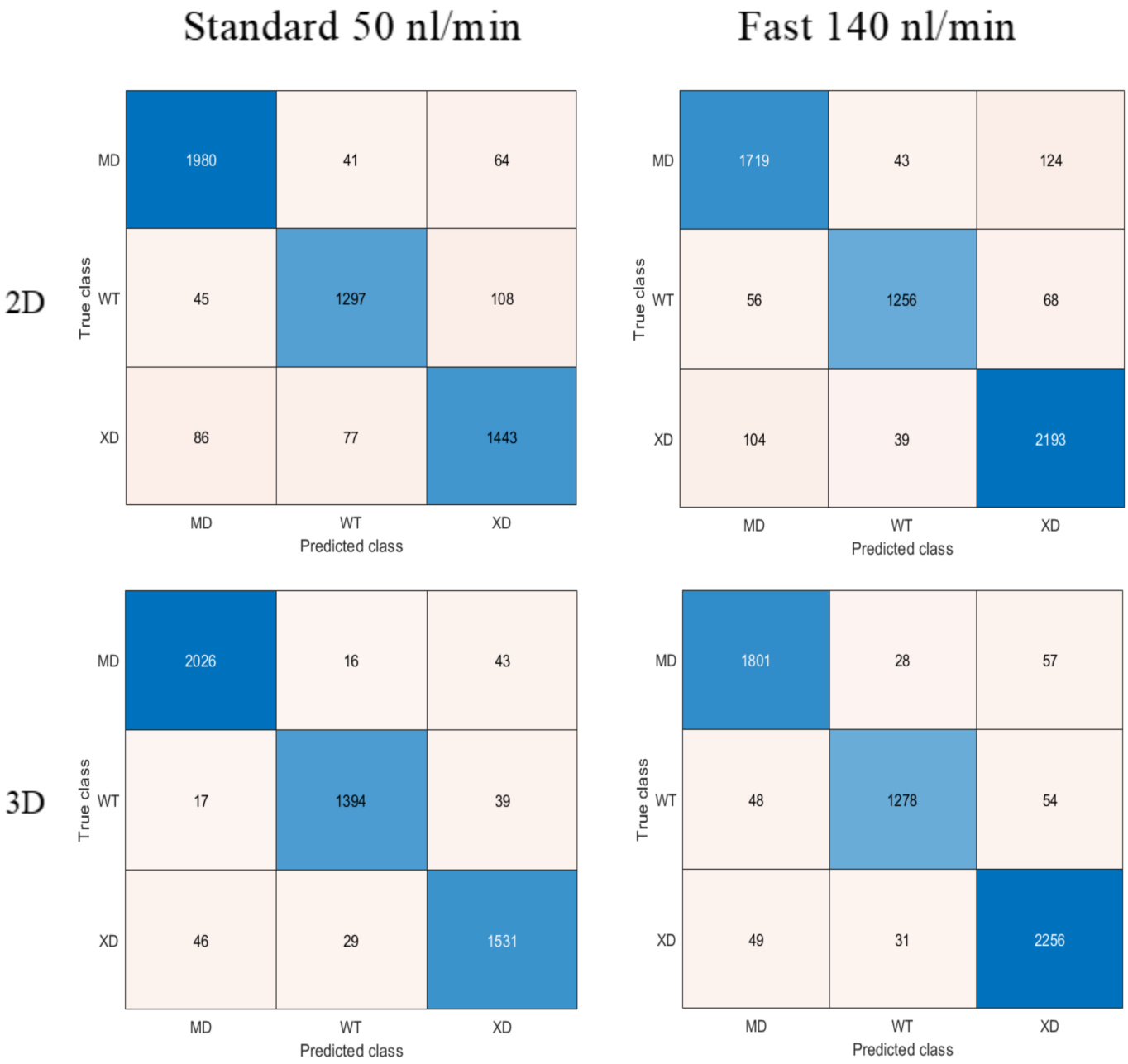

Fig. 8 Confusion matrices for the cells classification for standard and fast cell velocities based on 2D/3D scattering transform (SCATNET) features computed from $10-$folds cross validation method.

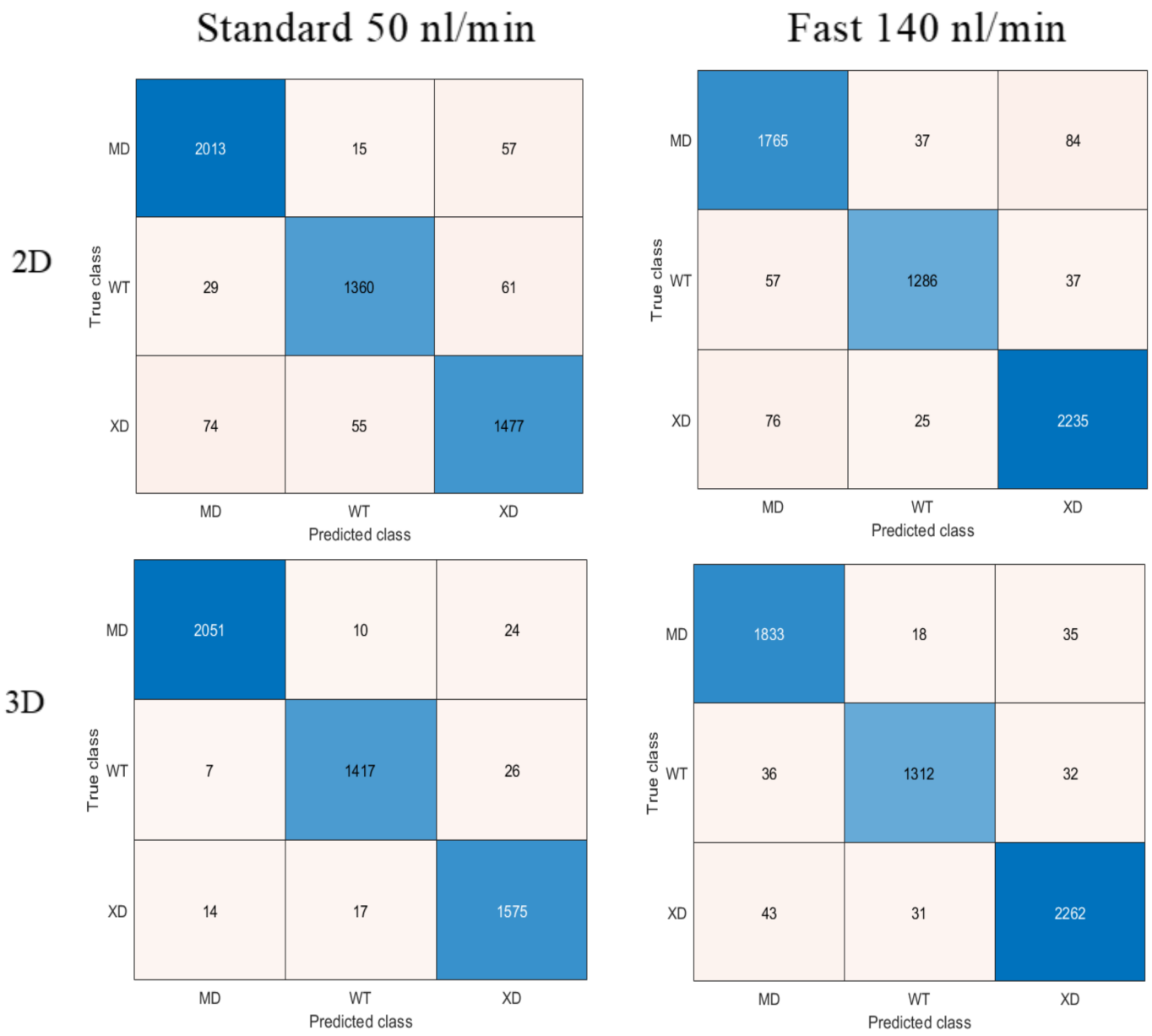

Fig. 9 Confusion matrices for the cell images classification for standard and fast cell velocities based on 2D/3D LBP, GLCM and SCATNET features (*All features*) concatenated together computed from 10−folds cross validation method.

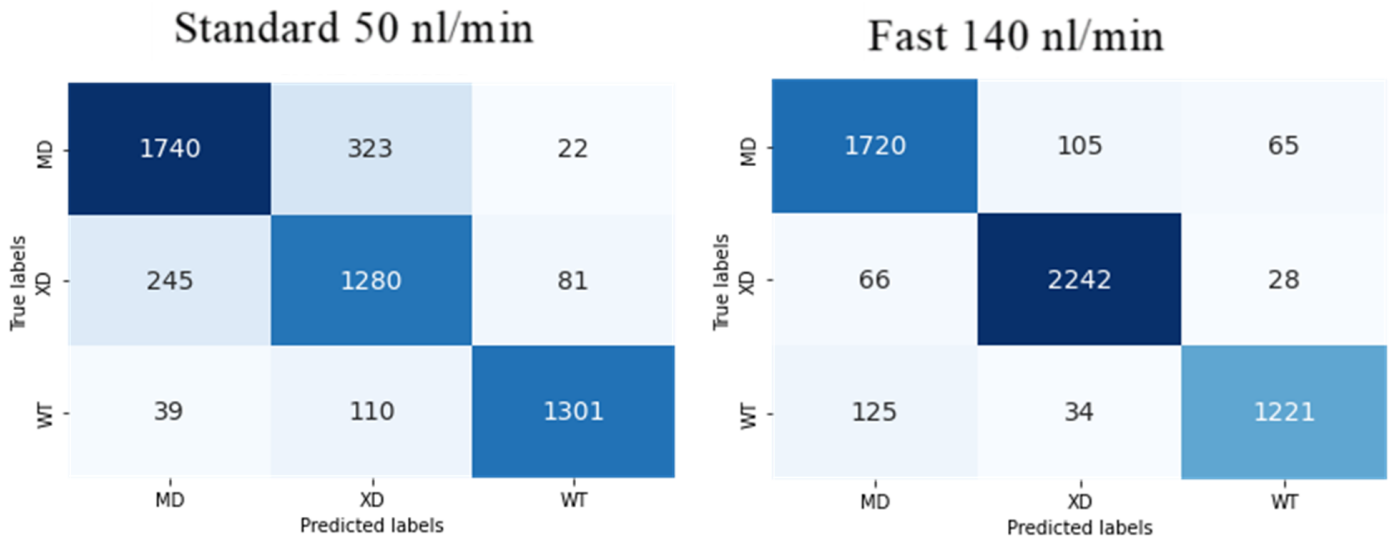

Fig. 10 Confusion matrices for the 2D cell images classification for standard and fast cell velocities based on convolutional neural network (CNN) computed from 10−folds cross validation method.

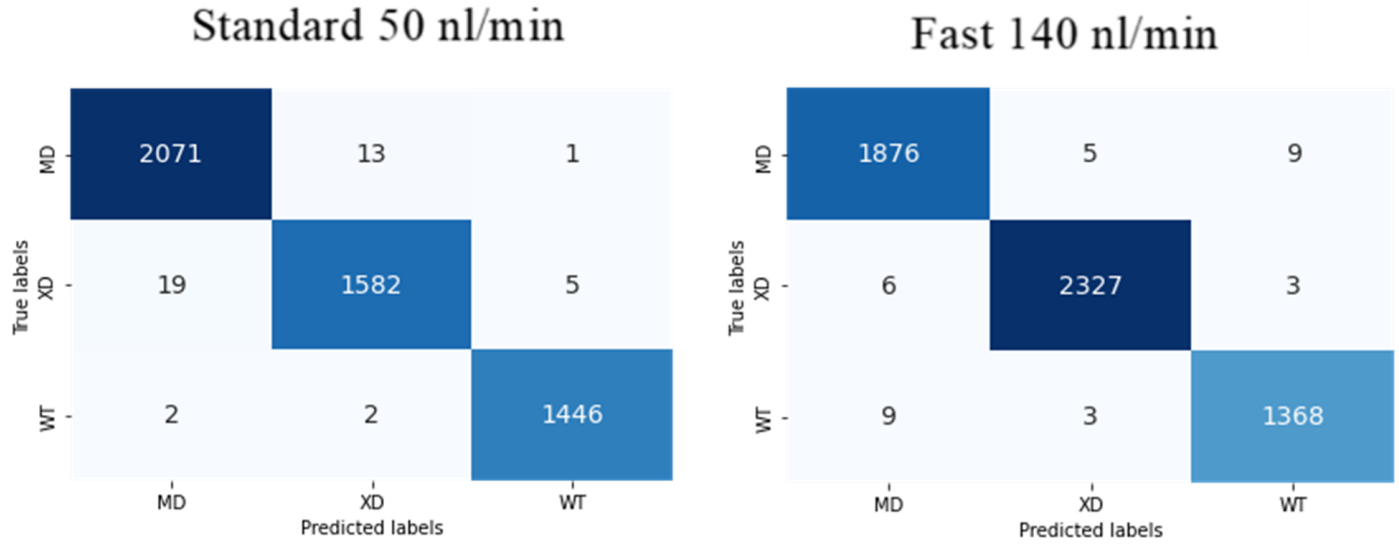

Fig. 11 Confusion matrices for the 2D cell images classification for standard and fast cell velocities based on convolutional neural network (CNN) with transfer learning from simulated images computed from 10−folds cross validation method.