Eurographics Workshop on Visual Computing for Biology and Medicine (2016)
H. Hauser, A. Hennemuth, and A. Lundervold (Editors)

Short Paper

A Feasibility Study on Automated Protein Aggregate Characterization Utilizing a Hybrid Classification Model

D. Eschweiler[1], M. Gadermayr[1], J. Unger[2], M. Nippold [3], B. Falkenburger[3], D. Merhof[1]

[1] ACTIVE Center, Institute of Imaging and Computer Vision, RWTH Aachen University
[2] Department of Biomedical Engineering, UC Davis
[3] Department for Neurology, Uniklinik RWTH Aachen

Abstract

The characterization of cytoplasmic protein aggregates based on time-lapse fluorescence microscopy imaging data is important for research in neuro-degenerative diseases such as Parkinson. As the manual assessment is time-consuming and subject to significant variability, incentive for the development of an objective automated system is provided. We propose and evaluate a pipeline consisting of cell-segmentation, tracking and classification of neurological cells. Focus is specifically on the novel and challenging classification task which is covered by relying on feature extraction followed by a hybrid classification approach incorporating a support vector machine focusing on mainly stationary information and a hidden Markov model to incorporate temporal context. Several image representations are experimentally evaluated to identify cell properties that are important for discrimination. Relying on the proposed approach, classification accuracies up to 80 % are reached. By extensively analyzing the outcomes, we discuss about strengths and weaknesses of our method as a quantitative assessment tool.

1. Motivation

Parkinson's disease (PD) is characterized by cytoplasmic aggregates of misfolded alpha-synuclein protein. For this reason, elucidating the formation and clearance of alpha-synuclein aggregates is a central effort to develop neuroprotective strategies for treatment of PD and potentially also for other neurogenerative diseases (e.g. Alzheimer's disease, amyotrophic lateral sclerosis). Research to develop novel therapies often starts at the cellular level. In cellular models, clearance of alpha-synuclein protein can be measured by different techniques, including immunoblots and microscopy [FSD16]. Aggregates are degraded by autophagy [DS16], in which parts of the cytosol are engulfed by a membrane, forming autophagosomes, which fuse with lysosomes for degradation. In this study, focus is on a frame-wise and cell-wise distinction between three cell-stages, namely normal, aggregate and dead. Visual characterization by medical experts so far relies on features such as fluorescence intensity distribution, geometrical appearance and movement characteristics. In order to study autophagic clearance of protein aggregates objectively, automated measurement tools that report protein aggregates are required.

1.1. Recent Work

Recently, significant research has been generally directed towards quantitative evaluation of microscopic imaging data [NHL*06, EBC*09, CZW06, VND*09, ZLYW09, CMT*16, SDFMM15]. Several application scenarios consist of cell phase identification [NHL*06, CZW06, ZLYW09, EBC*09, SDFMM15], DNA damage-repair pattern characterization [VND*09] and cancerous cell detection [CMT*16]. The analysis of protein aggregates, as considered in this work, is so far commonly performed biochemically or by means of flow cytometry [RPC*12, BOL*07]. Although imaging theoretically exhibits superior information content, automated assessment based on image analysis in this field, has not been investigated so far. As the output of microscopic imaging devices cannot be directly analyzed without further processing, image analysis methods are generally utilized for segmenting, tracking and finally classifying cells. For segmentation, various established image processing methods were deployed, including thresholding, morphological operations, region accumulation and deformable model fitting [Mei12]. Depending on the image data, tracking was performed relying on active-contours [PRTR09], methods relying on a frame-wise segmentation followed by graph matching [LZMW10] and probabilistic techniques [KFRJ06]. For cell classification, divergent approaches were recently proposed. In [EBC*09], an image representation consisting of an intensity histogram combined with a curvature histogram was introduced. Another approach aimed at a discrimination of several cell-stages based on a set of 23 statistical features (such as area, circularity) capturing both, shape and intensity information [VND*09]. Whereas these two approaches only incorporate spatial information, others also additionally exploit the temporal domain relying on heuristics concerning the sequence of states [CZW06] or on the

DOI: 10.2312/vcbm.20161277

Markov model [ZLYW09]. Finally, deep learning was recently applied to the task of cell classification [CMT*16]. Furthermore, for feature classification, established machine learning methods such as support vector machines (SVM), k-nearest neighbor classifiers and neural networks were utilized.

1.2. Contribution

We propose a pipeline for segmentation, tracking and classification of neurological cells in fluorescence image data. Whereas for segmentation and tracking established procedures from literature are utilized, focus is on the novel and challenging classification task. Specifically we discriminate for each cell and frame between three classes of fluorescence distribution: homogeneous (C1), aggregate (C2) and dead (C3) (Fig. 1). As this specific classification task has not been investigated so far, we analyze several different image representations to identify properties that are important for an effective discrimination. For feature classification, we deploy a hybrid model relying on a support vector machine (SVM) combined with a hidden Markov model (HMM). The SVM stage, which could be interpreted as dimensionality reduction, is inserted in order to improve robustness in combination with a relatively small set of training data. The HMM model is subsequently applied in order to additionally make use of the temporal context. By comparing the results of the SVM output with the outcomes of the hybrid SVM-HMM approach, a statement on the relevance of temporal context can be made.

2. Methods

The three stages of the proposed processing pipeline consisting of segmentation & tracking, feature extraction and classification, are described in detail in the following subsections.

2.1. Segmentation & Tracking

During image acquisition, five z-slices are captured with varying focus settings. Since cells are focused in different layers, first a maximum intensity projection is applied, providing the basis for the segmentation and tracking task. Segmentation is performed by thresholding each image slice by a predefined, fixed intensity value, which is based on the fact that images exhibit a homogeneous background [Mei12]. This leaves segmented objects including only cellular regions. Due to the low movement speed of cells, tracking is performed by registering centroids of segmented objects to the closest centroids in the subsequent frame. To increase robustness, a further constraint is introduced that cells in consecutive frames must overlap. If there is no overlap between cells, the objects are tagged as two individual cells. If two objects merge, the new object is tagged as consisting of multiple cells and is excluded from the evaluation stage to avoid any bias. Also, objects touching the image borders or being visible in less than 80 % of the sequence are labeled as unsuitable. This problem specific strategy was developed in cooperation and consensus with medical experts.

2.2. Feature Extraction

In order to extract most discriminative image representation, the feature extraction is limited to the cell's focus layer. The focus layers of individual cells are identified by determining the planes with the highest intensity value within the segmentation. Visual expert-based classification relies on several features, such as fluorescence intensity distribution, geometrical appearance and movement characteristics. Inspired by this policy and previous work [VND*09, NZ08], we identify five promising image representations, which are further analyzed and compared. The focus is specifically on examining conceptually different representations to obtain best possible insight about the features' performance.

1. Statistical static features:
   Strongly inspired by expert's classification as well as previous work [VND*09] in the field of cell classification, we identified a set of single (one dimensional) features extracting cell intensity, shape and texture properties. The image representation is formed by concatenating the following features:

   - Maximum intensity
   - Mean intensity
   - Entropy of intensity values
   - Standard deviation of intensity values
   - Pixel area of segmented object
   - Circularity calculated as quotient of minor axis length and major axis length $c_{object} = \frac{l_{minor}}{l_{major}}$.
   - Ratio of pixels with intensity equal to the local maximum intensity in the segmented object and the area of the segmented object $r_{max,object} = \frac{i_{max,object}}{a_{object}}$
   - Ratio of pixels with intensity values greater or equal to the global maximum intensity and the area of the segmented object $r_{max,global} = \frac{i_{max,global}}{a_{object}}$

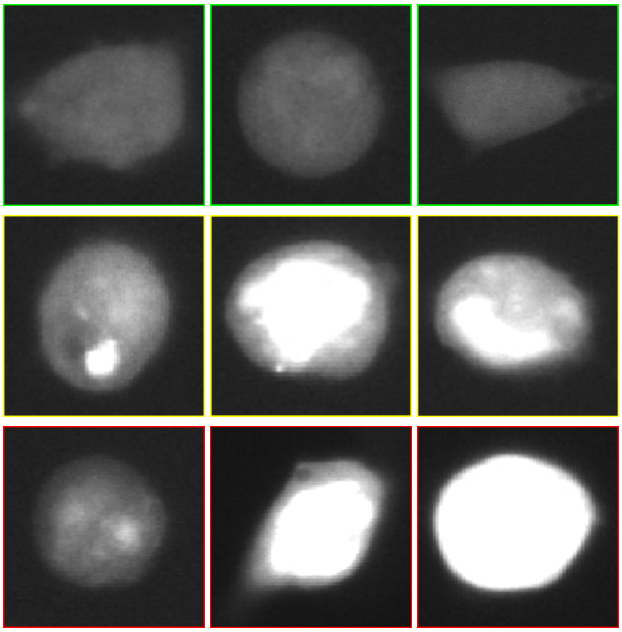

Figure 1: *Appearance of the three considered cell states (homogeneous (top row), aggregate (middle row), dead (bottom row)).*

The global maximum intensity $i_{max,global}$ referes to 95 % of the maximum intensity in the whole image sequence.

- Maximum aggregate correlation coefficient:
  Since the visual appearance of fluorescent proteins is characterized by their point-spread function, which can be adequately modeled by a 2D Gaussian [ZZOM07] and aggregates frequently appear as small bright globes, Gaussian distributions are used to locate aggregates within the segmented cells. The distributions comprise a mean value 0, standard deviation 1 and a radius varying from 1 to 5 pixels. For localization aggregate templates are created and the normalized correlation coefficient is used.

2. Statistical static & temporal features:
   To additionally incorporate temporal information (e.g. cell movement), this image representation contains all statistical static features (1.) as well as the following temporal ones:
   - Euclidean distance between centroids in two consecutive frames
   - Area change in two subsequent frames
     $a_{change,local} = a_{object,t} - a_{object,t-1}$
   - Area change referred to the mean value of all area changes within the image sequence
     $a_{change,global} = a_{object,t} - a_{mean}$
   - Circularity change in two subsequent frames
     $c_{change,local} = c_{object,t} - c_{object,t-1}$
   - Circularity change referred to the mean value of all ovality changes within the image sequence
     $c_{change,global} = c_{object,t} - c_{mean}$
3. Intensity histogram:
   The intensity histogram is computed based on a linear space and 16 bins. This method was already applied in previous work on cell classification [NZ08].
4. Local Binary Patterns [OPH94]:
   The well known and widely used Local Binary Patterns approach is deployed based on the standard eight pixel neighborhood in a circularly aligned pattern, resulting in a histogram consisting of 256 bins.
5. Feature fusion:
   Finally, all image representations (1. - 4.) are merged aiming at thereby increasing the discriminative power.

For classifier training, each single feature is normalized (zero mean and one standard deviation).

2.3. Classification

SVMs represent effective established classification models and were successfully applied in cell classification scenarios [NHL*06]. However, they are not capable of exploiting temporal context informations, which potentially leads to unsteady classification sequences in the investigated scenario. To overcome this problem, a hybrid SVM-HMM model as presented in [VP07] is utilized, which requires to further process the output obtained by the SVM. For the SVM stage, a one-vs-one classification procedure is preferred over a one-vs-all method, since it leads to more accurate predictions [VJIJ11]. The SVM raw-output provides distance measures between the feature points and the separating hyperplanes. Following Platt's method [Pla99] these distances are transformed to pairwise probabilities $p(q_i|\ or\ q_j, \mathbf{x})$ using a sigmoid function. q_i and q_j represent two classes of one classpair and $\mathbf{x}$ represents the feature vector. To obtain posterior probabilities $p(q_i|\mathbf{x})$ the following function is applied:

$$p(q_i|\mathbf{x}) = 1/\left[\sum_{j=1,j\neq i}^{3} \frac{1}{p(q_i|\ or\ q_j,\mathbf{x})} - 1\right] \quad (1)$$

Conditional probabilities $p(\mathbf{x}|q_i)$ are calculated by converting the posterior probabilities via Bayes' theorem,

$$p(\mathbf{x}|q_i) \propto \frac{p(q_i|\mathbf{x})}{p(q_i)} \quad (2)$$

where $p(q_i)$ are the a-priori probabilities for classes i which are attained from the training dataset. Due to proportionality, the emission probabilities have to be further normalized in order to set the sum over all classes equal to 1.

$$\sum_{i=1}^{3} p(\mathbf{x}|q_i) = 1 \quad (3)$$

Based on these extensions, the SVM stage delivers a three dimensional likelihood/probability vector, rather than a fixed label. The HMM model is then built on these likelihoods. Due to the low dimensionality and the ignorance of the distribution of the SVM output, we decided to utilize a discrete HMM model. To obtain discrete observations, vector quantization is applied. We specifically rely on k-means clustering, due to its combination of low cost intensity and effectiveness. To achieve the transition and emission probabilities for the HMM stage, a maximum likelihood estimation is performed on the training data set.

3. Experiments

3.1. Setup

For time-lapse microscopy we use an Olympus IX81 epifluorescence microscope (Olympus xcellence software, 40x air objective, NA 0.9) equipped with an incubator (37° C, 5 % CO2) and a motor stage to acquire images at defined positions every 30 minutes over 25 hours (50 frames). At each position, five z-slices are acquired with a z-distance of 2 μm to accommodate for differences in coverslip thickness. Fluorescence excitation is reduced to 3.95 % of maximum, exposure is 150 ms. For experimental purposes cells in 90 image sequences have been manually labeled by an expert neurologist, which provides a ground truth with a total of 136 cell sequences adding up to 3522 single cell images. To obtain conclusive and unbiased outcomes, we applied a policy of independently and repeatedly selecting 80 % of the image data for training, 10 % for parameter optimization and 10 % for evaluation. To further prevent bias, the training and the evaluation data is balanced in a way that the occurrence of all classes within the sets is equalized by means of random sampling. For the SVM, a linear kernel is used exhibiting good performance in similar application scenarios. Specifically, all combinations of the SVM's c-values ($2^{-1} - 2^{10}$) and the number of clusters ($2^2 - 2^{10}$) are computed and the optimized model is applied to the evaluation data set. Finally, the mean accuracy as well as the accuracy distribution of 200 iterations are reported. As the change between classes does not occur immediately, a classification

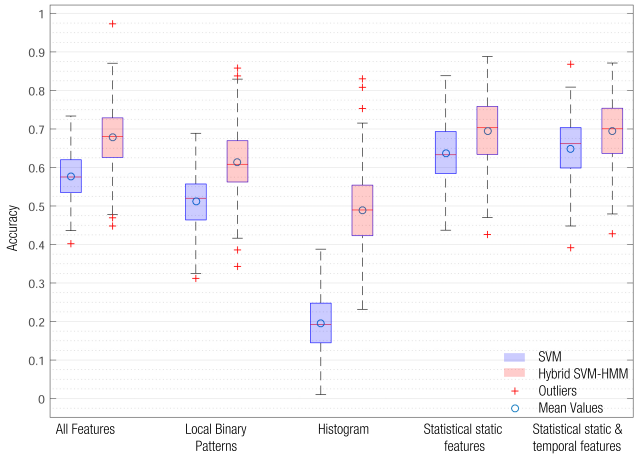

Figure 2: *Boxplots of accuracies for all feature compositions. Blue plots represent the SVM results and red plots represent the results of the hybrid SVM-HMM model.*

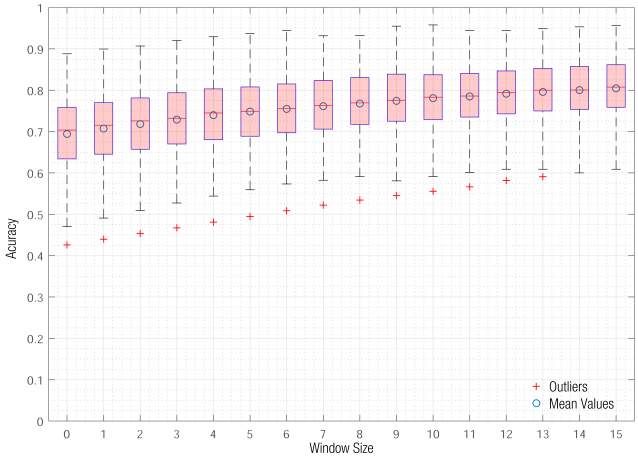

Figure 3: *Boxplots of prediction accuracies for increasing window sizes Δt using a SVM-HMM hybrid model for statistical static features*

of frames near state transitions often cannot be determined reliably even by manual inspection. To avoid bias due to vague data, analysis ignores windows around the transitions. The impact of variable window sizes $\Delta t = [1, 15]$ is furthermore analyzed.

3.2. Results

From manual expert segmentations we observed an ideal threshold of one-tenth the maximum intensity value (2^{12}) to exclude background and noise successfully. Low background noise prevents erroneous segmentations and sparse arrangement of cells allows strict exclusion of cells either touching each other or the image border without significantly reducing the data available for evaluation. To analyze the impact of the appended HMM stage and thereby the introduction of temporal context, we compare the SVM classification results to the hybrid model's performance. Results for SVM classification vary between a mean accuracy of 19.64 % for *Histograms* and 64.80 % for *statistical static features* (Fig. 2, blue plots). Furthermore, the confusion matrix is examined to get an insight of the classifiers strengths and weaknesses (Table 1). The classes C1 and C3 are predicted with an accuracy of 84.14 % and 87.06 % respectively, C2 impairs the overall results with an accuracy of 39.52 %. The hybrid SVM-HMM model is able to improve the mean accuracy for all feature compositions, particularly up to 69.48 % for *statistical static features* (Fig. 2, red plots). Moreover, the confusion matrix is corrected by balancing erroneous classifications of class 2 (Table 2). Hence, utilizing temporal context for classification sequences turns out to significantly and consistently improve the classification. Utilizing additional temporal, texture (Local Binary Patterns) or histogram features, however, does not further improve the results. Variable window sizes (as motived in Sect. 3.1) consistently increase the prediction accuracy further up to 80.39 % for a window size of 15 frames for *statistical static features* (Fig. 3). Considering this window size, class accuracies of 92.95 % (C1), 69.71 % (C2) and 82.01 % (C3) are obtained.

3.3. Discussion

Regarding the classification performances with the established SVM classification model combined with general purpose as well as domain specific features, the application scenario proved to be tough. Without considering the temporal context, overall classification rates of 64.80 % are obtained. Exploiting temporal context as well, the proposed hybrid HMM-SVM model improved the accuracy to 69.48 %. Obviously, the introduced knowledge on cell-state transmission can be effectively utilized to increase robustness. Considering the classification performance of the proposed method, we notice that the distinction between class C1 and the other classes can be performed quite precisely (87.64 % of class C1 samples were correctly classified). Difficulties were detected in case of a discrimination between class C2 and C3. 30.31 % of class C2 samples are categorized as class C3 and 24.64 % of class 3 samples are

Table 1: *Confusion matrix of SVM predictions trained by statistical static features*

Groundtruth \ Prediction	C1	C2	C3
C1	0.8414	0.0602	0.0984
C2	0.1472	0.3952	0.4576
C3	0.0368	0.0927	0.8706

Table 2: *Confusion matrix of SVM-HMM predictions trained by statistical static features*

Groundtruth \ Prediction	C1	C2	C3
C1	0.8764	0.0809	0.0427
C2	0.1612	0.5357	0.3031
C3	0.0392	0.2446	0.7162

wrongly categorized as class C2. This is obviously due to the high similarity between these two classes combined with a high intra-class variability (Fig. 1). Although the reported accuracies seem to be rather low, by increasing the window size, overall classification rates raised to 80.39 %. As the manual assessment is subject to inter-observer variability and state transitions are often vague, these rates are certainly more informative. To allow for a valid statistical analysis, especially the distinction between the classes C2 and C3 yields potential for improvements. The established image representation and classification methods investigated in this work show weaknesses in distinguishing between these two classes exhibiting a very high similarity. To increase reliability, in future work we will focus on two strategies: First, the amount of annotated training data will be increased to specifically obtain higher diversity. Additionally, annotations will be performed by multiple raters, to increase the ground truth's validity and to allow investigations of the inter-observer variability. As a second strategy, further image features will be developed specifically extracting markers with a high (visual) distinctiveness between the classes C2 and C3 derived from expert knowledge. As additional tool to increase the confidence of experimental outcomes, the certainty of each predicted sequence can be shown to the user along with the sequence labels. This allows to focus on the x % of most certain predictions and to drop cells that may have been wrongly classified. However, the inherent tradeoff between data reduction and precision improvement has to be considered.

4. Conclusion

In this paper, the aim was to automatically characterize fluorescent cells. To this end, we compared the SVM classifier with a hybrid SVM-HMM model. Cells were represented by various features, whose impact was further analyzed. SVM classification exhibited unsteady prediction sequences and thus inadequate results, since especially transitions are crucial. Appending a HMM classification stage enabled to exploit temporal informations and eliminated unsteady predictions. This resulted in a consistent improvement of the accuracy. It was further demonstrated, that the defined static features led to most accurate results.

Acknowledgments

This work was supported by the German Research Foundation (DFG), grant number ME3737/3-1.

References

[BOL*07] Basiji D. A., Ortyn W. E., Liang L., Venkatachalam V., Morrissey P.: Cellular image analysis and imaging by flow cytometry. *Clinics in Laboratory Medicine 27*, 3 (2007), 653–viii. 1

[CMT*16] Chen C. L., Mahjoubfar A., Tai L.-C., Blaby I. K., Huang A., Niazi K. R., Jalali B.: Deep learning in label-free cell classification. *Sci. Rep. 6* (Mar. 2016), 21471. 1, 2

[CZW06] Chen X., Zhou X., Wong S. T. C.: Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Transactions on Biomedical Engineering 53*, 4 (April 2006), 762–766. 1

[DS16] Dantuma N. P., Salomons F. A.: Ubiquitin versus misfolding: The minimal requirements for inclusion body formation. *J Cell Biol 213* (2016), 147–149. 1

[EBC*09] Ersoy I., Bunyak F., Chagin V., Cardoso M. C., Palaniappan K.: Segmentation and classification of cell cycle phases in fluorescence imaging. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'09)* (2009), pp. 617–624. 1

[FSD16] Falkenburger B. H., Saridaki T., Dinter E.: Cellular models for parkinson's disease. *Journal of neurochemistry* (2016), 1–14. 1

[KFRJ06] Kachouie N. N., Fieguth P., Ramunas J., Jervis E.: Probabilistic model-based cell tracking. *International Journal of Biomedical Imaging 2006* (2006), 1–10. 1

[LZMW10] Li F., Zhou X., Ma J., Wong S.: Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. *IEEE Transactions on Medical Imaging 29*, 1 (Jan. 2010), 96–105. 1

[Mei12] Meijering E.: Cell segmentation: 50 years down the road [life sciences]. *IEEE Signal Processing Magazine 29*, 5 (2012), 140–145. 1, 2

[NHL*06] Neumann B., Held M., Liebel U., Erfle H., Rogers P., Pepperkok R., Ellenberg J.: High-throughput RNAi screening by time-lapse imaging of live human cells. *Nature Methods 3*, 5 (May 2006), 385–390. 1, 3

[NZ08] Nilufar S., Zhang N. R. H.: Automatic blood cell classification based on joint histogram based feature and bhattacharya kernel. *42nd Asilomar Conference on Signals, Systems and Computers* (2008), 1915–1918. 2, 3

[OPH94] Ojala T., Pietikainen M., Harwood D.: Performance evaluation of texture measures with classification based on kullback discrimination of distributions. *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision and Image Processing, Proceedings of the 12th IAPR International Conference 1* (1994), 582–585. 3

[Pla99] Platt J. C.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in large margin classifiers* (1999), pp. 61–74. 3

[PRTR09] Padfield D., Rittscher J., Thomas N., Roysam B.: Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Medical Image Analysis 13*, 1 (2009), 143–155. 1

[RPC*12] Ramdzan Y. M., Polling S., Chia C. P. Z., Ng I. H. W., Ormsby A. R., Croft N. P., Purcell A. W., Bogoyevitch M. A., Ng D. C. H., Gleeson P. A., Hatters D. M.: Tracking protein aggregation and mislocalization in cells with flow cytometry. *Nature Methods 9*, 5 (Mar. 2012), 467–470. 1

[SDFMM15] Schönenberger F., Deutzmann A., Ferrando-May E., Merhof D.: Discrimination of cell cycle phases in PCNA-immunolabeled cells. *BMC Bioinformatics 16*, 1 (May 2015). 1

[VJIJ11] Varpa K., Joutsijoki H., Iltanen K., Juhola M.: Applying one-vs-one and one-vs-all classifiers in k-nearest neighbour method and support vector machines to an otoneurological multi-class problem. *Studies in health technology and informatics 169* (2011), 579–583. 3

[VND*09] Vos W. H. D., Neste L. V., Dieriks B., Joss G. H., Oostveldt P. V.: High content image cytometry in the context of subnuclear organization. *Cytometry 77*, 1 (2009), 64–75. 1, 2

[VP07] Valstar M., Pantic M.: Combined support vector machines and hidden markov models for modeling facial action temporal dynamics. *Human Computer Interaction 4796* (2007), 118–127. 3

[ZLYW09] Zhou X., Li F., Yan J., Wong S.: A novel cell segmentation method and cell phase identification using markov model. *IEEE Transactions on Information Technology in Biomedicine 13*, 2 (Mar. 2009), 152–157. 1, 2

[ZZOM07] Zhang B., Zerubia J., Olivo-Marin J.: Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt. 46*, 10 (2007), 1819–1829. 3