A Comparative Study of Machine Learning Algorithms and SIFT for Identifying the Quality of Human Induced Pluripotent Stem Cell Colony

Henry Joutsijoki
Faculty of Information Technology and Communication Sciences
Tampere University
Tampere, Finland
henry.joutsijoki@tuni.fi

Markus Haponen, Katriina Aalto-Setälä
Faculty of Medicine and Health Technology
Tampere University
Tampere, Finland
markus.haponen@tuni.fi, katriina.aalto-setala@tuni.fi

Martti Juhola
Faculty of Information Technology and Communication Sciences
Tampere University
Tampere, Finland
martti.juhola@tuni.fi

***Abstract*—Personalized medicine took giant steps further when the human induced pluripotent stem cell (hiPSC) technology was introduced. hiPSC technology enables reprogramming human somatic cells back to stem cells and differentiate them into any cell type wanted, and use them, for instance, in treatments and pharmacological research. Growing and differentiating hiPSCs is a sensitive process that requires constant monitoring since during the growth cycle unwanted changes in hiPSC colonies may occur which prevents their further use. To automate the monitoring and the quality control of hiPSC colonies, we need computational tools. In this paper, we examine the quality identification of hiPSC colony images by comparing a collection of machine learning methods and applying Scaled Invariant Feature Transformation (SIFT) algorithm in classification. Our dataset consists of altogether 229 hiPSC colony images from good/semigood/bad classes and having three quality classes separates our study from the other researches in this field. We obtained 73.4% accuracy using K-Nearest Neighbor (KNN) classifier which outperforms the results from our earlier researches with above 10%. The results show that the quality control of hiPSC colonies can be performed with good accuracy and with methods that are transparent and suitable in medical domain.**

***Index Terms*—Machine learning, SIFT, Human induced pluripotent stem cell, Personalized medicine**

I. Introduction

The reprogramming of somatic cells back to stem cells was a revolutionary invention by Takahashi et al. [26], [27] and changed the course of personalized medicine. The produced cells with the introduced method were called human induced pluripotent stem cells (hiPSCs) and the technique enabled a non-invasive way to differentiate into cell types wanted such as cardiomyocytes or nerve cells. hiPSC technique gave a totally new direction for the stem cell research and it has been extensively applied in different medical applications, such as in drug research [13], [18].

Rapid increase and utilization of hiPSCs both in research and clinical use has led to a situation where researchers and practitioners need decision supportive tools for their daily work. Nowadays, the use of artificial intelligence (AI) and machine learning (ML) has gained a firm foothold in many domains and applications, and utilizing the potential of AI and ML has decreased the amount of manual routine-based work. This again has released resources to more complex tasks from the human experts. In hiPSC research, one focal task is to monitor the growing stem cell colonies to track possible abnormalities and exclude unsuitable colonies from the further examination.

In our earlier research, we applied histogram-based [6], [11] features to identify the status of hiPSC colony image whereas in [7], [8] we evaluated the average SIFT [20], [21] descriptor image-wise and used it in classification. From the classification method point of view, a collection of off-the-shelf classification algorithms likewise Support Vector Machine (SVM) variants were used. The current trend in machine learning is that deep learning solutions are applied to almost every task as a baseline method. When it comes to iPSC research, there is a similar trend ongoing and deep learning methods, especially Convolutional Neural Networks (CNNs), are utilized in computational analysis of iPSC colony images.

Fan et al. [4] analyzed iPSCs colony images with HMM (Hidden Markov Models) to model the growth curve of iPSCs, CNN for the segmentation of images to classify the content of an iPSC colony image into non-iPS and iPS parts. Kavitha et al. [14] examined classification and segmentation of iPSC colony images (healthy/unhealthy). Segmentation was performed by adapting k-means algorithm and in the classification SVM, random forest, multi-layer perceptron, decision tree and Adaboost classifiers were applied. From the images, shape, moment, statistical, spectral features, and their combinations were examined and an accuracy above 90% was achieved. Kavitha et al. [15] proposed a vector-based Convolutional

979-8-3503-1017-7/23/$31.00

Neural Network (V-CNN) approach for automatic recognition of iPSC colonies (healthy/unhealthy) and it outperformed the SVM results having accuracies above 90%. Piotrowski et al. [24] applied a convolution based encoder-decoder based on U-Net solution to segment and classify hiPSC culture status whereas Orita et al. [23] proposed to use VGG16-based CNN to classify hiPSC colonies into normal/abnormal classes. Orita et al. [23] achieved F1 score of 0.89. Yue et al. [32] introduced a combination of CNN and SVM approach to evaluate iPSC colony quality and achieved around 95% accuracy. Besides the aforementioned studies, Issa et al. [5] and Coronnello and Francipane [3] conducted reviews on how ML/AI methods have been used in the analysis of stem cell derived data. All these studies show how important and timely topic the analysis of hiPSC colony images is by means of machine learning and computer vision methods.

In this paper, our aim is to show that "traditional" computer vision and classification methods have still a place in applied machine learning research and, especially, in the context of computational analysis of hiPSC colony images. We use SIFT descriptors and a collection machine learning methods to classify hiPSC colony images into one of the three classes (good/semigood/bad). Dividing the hiPSC colonies into three classes separates our study from the other researches in this application domain and gives a unique aspect to this paper. Compared to our earlier studies regarding the classification of hiPSC colony images [6]–[8], [11], we have several differences in this paper that are summarized as follows:

1) Larger image dataset
2) Different way of utilizing SIFT descriptors
3) Different preprocessing of hiPSC colony images
4) New classification methods.

The rest of the paper is organized as follows. Section II describes the dataset and classification methods used as well as preprocessing and classification set-up. In Section III, we present the classification results and Section IV concludes the paper.

II. Methods and experiments

The study was approved by the ethical committee of Pirkanmaa Hospital District (R08070). Establishing of iPSC lines was performed with retroviruses encoding for OCT4, SOX2, KLF4, and MYC as described in detail in [26]. Cell line characterization for their karyotypes and pluripotency has been presented in [17].

A. Data acquisition

The image acquisition process in this paper followed the same procedures as given in our earlier publications [6]–[8], [11]. hiPSCs were used in this paper and the colonies were imaged between days 5-7 of their weekly growth cycle [6]–[8], [11]. The rationale for selecting days 5-7 is that during these days better visualization of hiPSC colonies is obtained. The growing hiPSC colonies were observed before taking the image from the colony and after observation labeling the image into one of the classes (good/semigood/bad) was performed [6]–[8], [11]. The imaging equipment used in this study was Nikon Eclipse TS100 inverted routine microscope with an attached heating plate [6]–[8], [11]. Furthermore, image acquisition equipment was performed using Imperx IGV-B1620M-KC000 camera which was mounted to the microscope and connected to a notebook equipped with JAI Camera Control Tool software [6]–[8], [11].

During the imaging process, lighting and sharpness of an image were manually defined which may produce some differences between images [6]–[8], [11]. However, the same human expert performed the data acquisition process in this study to minimize the variability between different human experts. Moreover, imaging settings were fixed during the photographic sessions but the image data was obtained in several sessions which causes some minor variability in the images [6]–[8], [11]. In majority of the cases, hiPSC colonies were in the center of the image, which gives the best visual condition [6]–[8], [11]. Nevertheless, in some cases observed colony was near the edge of the well and this caused some distortion in the lightning [6]–[8], [11].

B. Dataset

The image dataset includes images from three classes (good/semigood/bad) and the categorization of the hiPSC colonies followed the same rules as in our earlier researches [6]–[8], [11]. The definitions for each class are summarized as follows [6]–[8], [11]:

1) Label good is given when hiPSC colonies had rounded shape, translucent even color, and defined edges.
2) Label semigood is given for hiPSC colonies that present changes in color and structure, but still had clear edges.
3) Label bad is given for the hiPSC colonies that had partially lost the edge structure, vacuole could occasionally be observed and areas of three-dimensional structures were observed.

Besides the actual hiPSC colony images, a human expert has manually constructed a mask for each colony image that has been utilized in preprocessing stage. Furthermore, the same human expert labelled the images and the labeling was used in classification as ground truth information. Table I shows the class sizes in the dataset and the corresponding proportion. As shown in Table I, class good is the largest class and class bad is the smallest class in a dataset. Figure 1 presents two example images from each class with corresponding masks. Resolution for all images was 1608×1208 (width$\times$height).

Table I
DATASET DESCRIPTION.

Class	Class size	Proportion (%)
Bad	58	25.3%
Good	94	41.0%
Semigood	77	33.6%
Total	229	100%

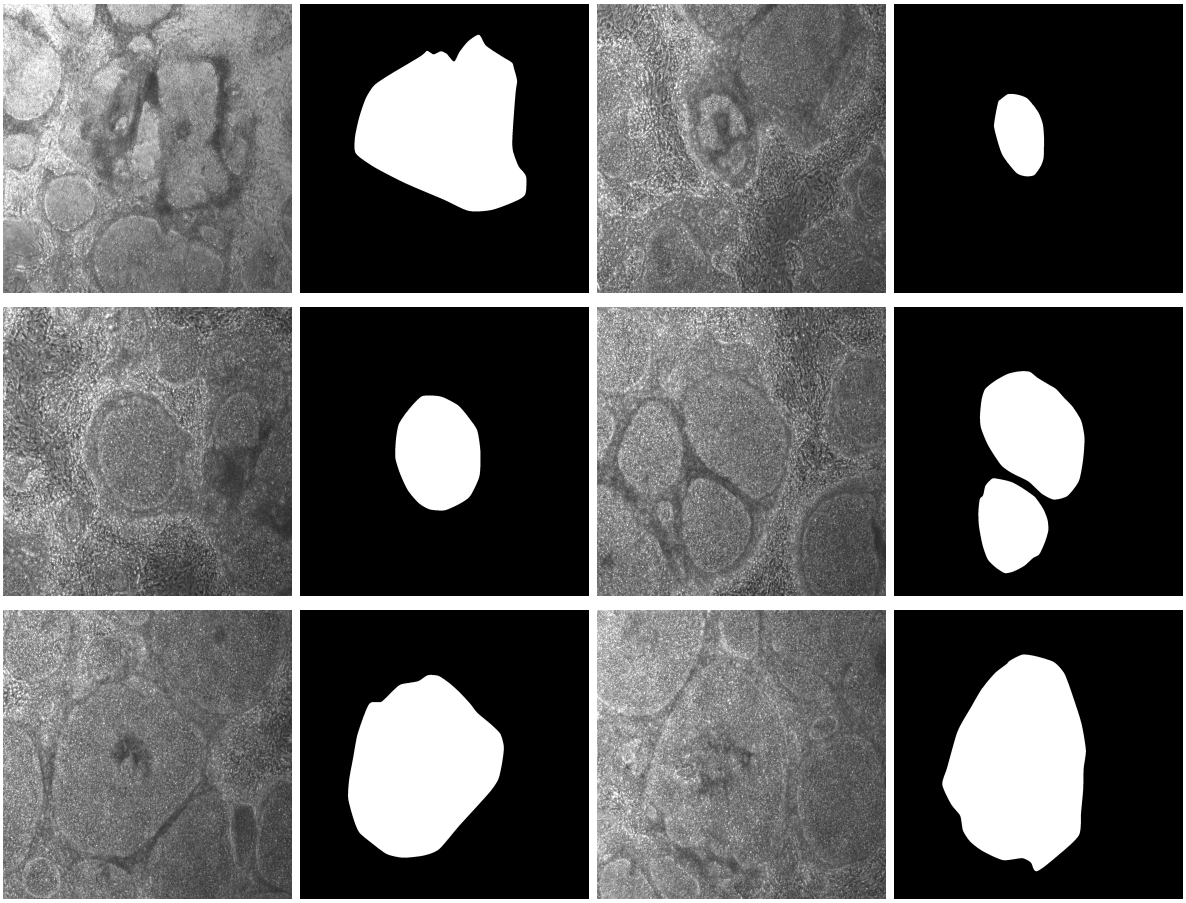

Figure 1. Example images of hiPSC colonies. Top row represents bad colony images and their masks. Middle row presents good colony images and the bottom row illustrates semigood colony images. Images have been scaled to have width and height of 1.5in.

C. Classification methods

In this study, we applied altogether seven different machine learning methods. Some of the methods have been used also in our earlier researches [6]–[8], [11], but to this paper we have included new classification methods. The first method is K-Nearest Neighbor (KNN) method [1], which is a distance-based classifier and known to be one of the top 10 data mining algorithms [31]. The performance of KNN classifier is dependent on three main parameters: distance measure, distance weighting, and the selection of k value. We did not vary the distance weighting and kept it equal for all instances in a dataset. We tested five different distance measures which were Chebyshev, cityblock (also known as Manhattan distance), cosine, Euclidean, and Mahalanobis. In the case of each distance measure, we examined 8 values of k ($k \in \{1, 3, 5, 7, 9, 11, 13, 15\}$) and we chose to use only odd values since it decreases the possibility of having a tie in KNN classification. The rationale of including KNN into this study is that it performed best, for instance, in [7], [8]. Moreover, with KNN it is possible to have a detailed information about the hiPSC colonies which are the most important ones related to the hiPSC colony image to be classified.

Two discriminant analysis methods were used in this paper which were Linear Discriminant Analysis (LDA) [29] and Quadratic Discriminant Analysis (QDA) [28]. Discriminant analysis methods are, generally speaking, good baseline methods for classification tasks and they have been applied in many iPSC researches (see, for example, [12], [13]). Another parameter free method applied in this research is multinomial logistic regression (MLR) [16], an extended version of two-class logistic regression method. MLR has been used in other image classification tasks successfully, such as in the classification of benthic macroinvertebrate images [9] which was the motivation to apply it also in this study.

From the family of artificial neural networks methods, we chose Probabilistic Neural Networks (PNN) [25] to be used in this study. PNN has been used in benthic macroinvertebrate image classification [10] with good results. PNN has a relatively simple architecture compared to other neural network methods which attracts the practitioners to use it in novel domains and applications. PNN requires tuning a hyperparameter of σ that is the width of Radial Basis Function and we tested the values of $\sigma \in \{0.125, 0.25, 0.5, 1, 2, 4\}$. Two tree-based classification methods were used in our study and these were Classification and Regression Trees (CART) [19] and random forests (RF) [2]. Overall, tree-based methods are transparent

and the model obtained is easy to understand. Moreover, tree-based methods are computationally efficient. Furthermore, CART and RF have given in other hiPSC applications [12], [13] very good results which was the motivation for using them in hiPSC colony image classification. Since the performance of RF classifier depends highly on the selection of the number of trees, we tested 20 different settings for RF classifier ($\#trees \in \{5, 10, 15, \ldots, 100\}$).

D. Preprocessing

Preprocessing for the hiPSC colony images was a multi-staged process. In the experimental part of the paper, we have two research lines what we examined in a more detailed way. These research lines differ from the preprocessing point of view. The complete preprocessing procedure can be described as follows:

1) Perform contrast-limited adaptive histogram equalization (CLAHE) [22] for the whole hiPSC colony image.
2) Perform segmentation for the hiPSC colony image by applying the predefined masks.
3) Extract SIFT [20], [21] descriptors from the segmented hiPSC colony image.

The first research line includes all the presented preprocessing stages and is called histogram equalized hiPSC colony images. The other research line is otherwise the same as the first one, but does not include CLAHE stage for the hiPSC colony images and is called histogram non-equalized hiPSC colony images.

Extraction of SIFT descriptors was performed by using VLFeat [30] and the default values of SIFT given by the VLFeat was used. An in-depth presentation of SIFT algorithm has been presented in [20], [21] and in [8] the main points of the method have been given. The main idea of SIFT is to find keypoints from the image and present them with 128-dimensional vectors called descriptors. The extraction process of SIFT descriptors includes four stages which are scale-space extrema detection, keypoint localization, orientation assignment, and computing the keypoint descriptors [8], [20]. In the case of the first research line, the total number of SIFT descriptors was around 778K. In the second research line, the total number SIFT descriptors was around 649K.

E. Classification procedure and performance measures

We performed classification using leave-one image data-out approach and the classification process can be described as follows:

1) Exclude the SIFT descriptors of the ith image to a test set.
2) Construct a training set from the rest of the SIFT descriptors.
3) Perform z-score standardization for the training set.
4) Standardize the test set using the scaling parameters gained from the training set.
5) Train algorithm using training set and predict class labels for the standardized test set from the ith image.
6) Take the mode of the predictions for the test set SIFT descriptors as a final predicted class label for the ith image.
7) Repeat stages 1-6 for all hiPSC colony images.

If a classification method required parameter tuning, the presented classification process was repeated with all parameter values tested. The best parameter value was selected based on the highest accuracy. In the experimental part of the paper, we have presented in the result tables the best parameter values for parameter dependent classification methods.

We selected three standard performance measures to be used in this paper. The rationale for selecting these performance measures is that they are directly applicable to multi-class classification task. Many of the performance measures (e.g., F1-score) are originally designed for binary classification tasks, and extending them to multi-class case requires evaluating and presenting the performance measure for each class separately, for example, using one-vs.-rest approach.

The performance measures applied in this study are true positive (TP), true positive rate (TPR) also known as sensitivity/recall, and accuracy. All performance measures were evaluated from a confusion matrix. TP stands for the number of correctly classified samples from a specific class, whereas TPR is the proportion of correctly classified samples from a specific class. Accuracy is defined by summing up the number of correctly classified samples and dividing it by the sum of all elements in a confusion matrix.

III. Results

Table II shows the results of hiPSC colony image classification when the images were histogram non-equalized in the preprocessing phase. When looking first the KNN results in a more detailed way, we notice that there is a spread in terms of accuracies since they vary from around 54% to nearly 69%. This notice alone strengthens the basic assumption that a thorough examination of different distance measures is a mandatory task to perform to find optimal settings for the task at hand.

The highest accuracy, 68.6%, was gained with the KNN classifier using cityblock metric whereas the lowest KNN-based accuracy was obtained with the Chebyshev metric. A common detail within all KNN results is that small k values were the best ones which may indicate clear locality with respect to the SIFT descriptors extracted from the images. Furthermore, when the k value is large, more noise is involved which again may decrease the reliability of classification. The closer inspection of TPs and TPRs reveal that the class good was recognized consistently with a good level using KNN. TPs were with KNN 70 or over and TPRs at least 74.5% respectively. One reason behind the accurate identification of good colony images may be that it was the largest class in the dataset having 94 images from 229 images altogether. Another reason might be that the variability within the good colony images can be smaller indicating more homogeneous class and this leads to a situation where the SIFT descriptors from good images are not dispersed so much in a data space.

Table II
CLASSIFICATION RESULTS OF HISTOGRAM NON-EQUALIZED HIPSC COLONY IMAGES.

	True positive			True positive rate (%)			
Method	Bad	Good	Semigood	Bad	Good	Semigood	Accuracy (%)
KNN Chebyshev ($k = 3$)	25	89	10	43.1	94.7	13.0	54.1
KNN cityblock ($k = 1$)	40	82	35	69.0	87.2	45.5	**68.6**
KNN cosine ($k = 1$)	55	70	26	94.8	74.5	33.8	65.9
KNN Euclidean ($k = 1$)	20	89	28	34.5	94.7	36.4	59.8
KNN Mahalanobis ($k = 1$)	15	87	28	25.9	92.6	36.4	56.8
CART	52	77	18	89.7	81.9	23.4	64.2
LDA	46	91	0	79.3	96.8	0.0	59.8
QDA	41	92	0	70.7	97.9	0.0	58.1
MLR	45	91	0	77.6	96.8	0.0	59.4
PNN ($\sigma = 0.5$)	20	88	28	34.5	93.6	36.4	59.4
Random forests ($\#trees = 5$)	56	92	4	96.6	97.9	5.2	66.4

In the case of bad class, TPR and TP values vary greatly between different distance measures. The lowest TPR, 25.9%, was obtained with the Mahalanobis metric and the highest TPR and TP values were gained with cosine measure. This shows how sensitive the classification may be with different distance measures and bad hiPSC colony images can include more diverse images compared to good images that can have an influence to the results. Semigood class can be considered as a transition phase from good to bad. Table II results show that semigood class was the most difficult class to recognize. TP values differ from 10 to 35 and TPRs 13.0% to 45.5% respectively with KNN. The difficulty of classifying semigood class is understandable since some of the images can be more close to bad images than good images and vice versa. Hence, there can be easily confusion to other classes in the case of semigood images.

When considering the results from other methods, topmost results are left behind the best KNN result. RF and CART, both tree-based methods, achieved above 64.0% accuracy whereas the rest of the methods yielded below 60.0% accuracy. Accuracies of CART and RF still outperformed the best accuracies from our earlier researches [6]–[8], [11]. The same tendency regarding the semigood class continued what was visible with KNN results, and the highest TPR was 36.4% gained by the PNN. Good colonies were identified well with other methods than KNN, and with the bad hiPSC colony images recognition rate was also good except with the PNN. Overall, the success of KNN is not a big surprise since in [7], [8] KNN was the best method among all methods tested.

Table III presents the results of hiPSC colony image classification with SIFT descriptors and when the images are histogram equalized. The general trend of results is similar to Table II results but there are still some differences. The accuracies gained by the KNN variants are lower compared to Table II except KNN with cosine measure that obtained an accuracy of 73.4%. This accuracy is the first time when a limit of 70.0% was surpassed in hiPSC colony image classification when taking into account our previous researches [6]–[8], [11]. Otherwise, accuracies obtained by the KNN variants were left below 70.0% and were worse than the corresponding accuracy in Table II. An interesting detail in KNN cosine result is that bad colony images were recognized better than the good ones. Usually, the situation has been the opposite. Furthermore, semigood hiPSC colony images were classified with above 60.0% recognition rate, and this is clearly higher result compared to other methods in Tables II-III. Histogram equalization and standardization as a preprocessing steps may have had the necessary influence to receive such a good accuracy. With cityblock and Euclidean metrics, KNN achieved above 40.0% TPR on semigood class. Good hiPSC colony images were identified above 70.0% TPR with all KNN variants and in the case of bad hiPSC colony images KNN with cosine and Mahalanobis distance measures above 70.0% TPR was gained. When examining TP values in Table III the same phenomena can be seen as for TPRs.

For other methods than KNN, the accuracies in Table III are better compared to Table II results except for PNN. Now, five methods reached above 60.0% accuracy whereas only two methods got this kind of result in Table II. CART and RF were better than LDA, QDA, MLR, and PNN and this follows the similar trend than in Table II. Outside KNN methods, the semigood class was again the most difficult class to identify. The TPRs ranged from 0.0% to 40.3% and TPs from 0 to 31 respectively. Good hiPSC colony images were recognized with at least 85.0% TPR, which is a very good result. Moreover, bad colony images were classified with over 82.0% TPR except in the case of PNN. Overall, we can say that the general level of results with histogram equalized hiPSC colony images was better than histogram non-equalized images. KNN was the best method as it was in our earlier publications [7], [8].

IV. CONCLUSIONS

This paper dealt with the classification of hiPSC colony images which has received in recent years growing interest in scientific community. The rationale for the interest towards hiPSC colony image classification comes from the fact that personalized medicine has taken great leaps towards in developing individual treatments and stem cells play a key role in this domain. In the future while conducting large-scale cell culture studies, the manual quality control process is not

Table III
CLASSIFICATION RESULTS OF HISTOGRAM EQUALIZED HIPSC COLONY IMAGES.

	True positive			True positive rate (%)			
Method	Bad	Good	Semigood	Bad	Good	Semigood	Accuracy (%)
KNN Chebyshev ($k = 3$)	17	86	8	29.3	91.5	10.4	48.5
KNN cityblock ($k = 1$)	32	79	38	55.2	84.0	49.4	65.1
KNN cosine ($k = 1$)	54	67	47	93.1	71.3	61.0	**73.4**
KNN Euclidean ($k = 1$)	12	89	31	20.7	94.7	40.3	57.6
KNN Mahalanobis ($k = 3$)	49	72	9	84.5	76.6	11.7	56.8
CART	48	80	28	82.8	85.1	36.4	68.1
LDA	52	91	0	89.7	96.8	0.0	62.4
QDA	53	91	0	91.4	96.8	0.0	62.9
MLR	52	91	0	89.7	96.8	0.0	62.4
PNN ($\sigma = 0.5$)	12	90	31	20.7	95.7	40.3	58.1
Random Forest ($\#trees = 5$)	54	92	12	93.1	97.9	15.6	69.0

feasible anymore and human experts need supportive tools for excluding unsuitable hiPSC colonies to ensure the best possible treatment and safety for patients.

We introduced a SIFT-based solution to the classification task that differed from our earlier researches [7], [8] where also SIFT were used. Besides the use of SIFT, we examined different preprocessing approaches to find out the optimal settings for the classification task. Our dataset included three classes (good, bad, and semigood) and this gives a unique aspect compared to other studies examining the hiPSC colony quality assessment where the problem has been reduced to two class classification task (good/healthy vs. bad/unhealthy). However, the semigood class enables earlier identification of unwanted changes in hiPSC colonies. Since the growing process hiPSC colonies is basically a time-series task, the semigood class presents the transition phase from good to bad.

We examined a group of seven machine learning methods consisting of methods (PNN and RF) that have not been used in our earlier researches [6]–[8], [11]. The best result was accuracy of 73.4% obtained by the KNN with the cosine distance measure and the success of KNN classifier is in line with our earlier publications. The best result also outperformed our earlier topmost results with over 10% and showed that the hiPSC colony image classification to good, bad, and semigood classes is possible with good accuracy. The result also strengthens our approach to use three classes instead of simplified good vs. bad classification.

Although we have applied "classical" computer vision and machine learning methods, our results are highly promising. The methods used in this study have still their own place in applied machine learning research and should not be forgotten even though deep learning has become the leading trend in majority of the applications. Deep learning has a firm foothold in machine learning research currently and in future our aim is to move towards this approach, and to examine whether or not they will outperform "traditional" computer vision and machine learning methods.

REFERENCES

[1] H.A.A. Alfeilat, A.B.A. Hassanat, O. Lasassmeh, A.S. Tarawneh, M.B. Alhasanat, H.S.E. Salman, and V.B.S. Prasath, "Effects of distance measure choice on KNN classifier performance - A Review," *Big Data*, vol. 7, no. 4, pp. 221-248, 2019.

[2] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

[3] C. Coronnello and M.G. Francipane, "Moving towards induced pluripotent stem cell-based therapies with artificial intelligence and machine learning," *Stem Cell Reviews and Reports*, vol. 18, pp. 559-569, 2022.

[4] K. Fan, S. Zhang, Y. Zhang, J. Lu, M. Holcombe, and X. Zhang, "A machine learning assisted, label-free, non-invasive approach for somatic reprogramming in induced pluripotent stem cell colony formation detection and prediction," *Scientific Reports*, vol. 7, Article number 13496, 2017.

[5] J. Issa, M.A. Chaar, B. Kempisty, L. Gasiorowski, R. Olszewski, P. Mozdziak, and M. Dyszkiewicz-Konwinska, "Artificial-intelligence-based imaging analysis of stem cells: A systematic scoping review," *Biology*, vol. 11, no. 10, Article 1412, 2022.

[6] H. Joutsijoki, M. Haponen, I. Baldin, J. Rasku, Y. Gizatdinova, M. Paci, J. Hyttinen, K. Aalto-Setälä, and M. Juhola, "Histogram-based classification of iPSC colony images using machine learning methods," in *Proceedings of the 2014 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2611-2617, 2014.

[7] H. Joutsijoki, M. Haponen, J. Rasku, K. Aalto-Setälä, and M. Juhola, "Error-correcting output codes in classification of human induced pluripotent stem cell colony images," *BioMed Research International*, vol. 2016, Article ID 3025057, pp. 1-14, 2016.

[8] H. Joutsijoki, M. Haponen, J. Rasku, K. Aalto-Setälä, and M. Juhola, "Machine learning approach to automated quality identification of human induced pluripotent stem cell colony images," *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 3091039, pp. 1-16, 2016.

[9] H. Joutsijoki and M. Juhola, "A comparison of classification methods in automated taxa identification of benthic macroinvertebrates," *International Journal of Data Science*, vol. 2, no. 4, pp. 273-300, 2017.

[10] H. Joutsijoki, K. Meissner, M. Gabbouj, S. Kiranyaz, J. Raitoharju, J. Ärje, S. Kärkkäinen, V. Tirronen, T. Turpeinen, and M. Juhola, "Evaluating the performance of artificial neural networks for the classification of freshwater benthic macroinvertebrates," *Ecological Informatics*, vol. 20, pp. 1-12, 2014.

[11] H. Joutsijoki, J. Rasku, M. Haponen, I. Baldin, Y. Gizatdinova, M. Paci, J. Saarikoski, K. Varpa, H. Siirtola, J. Avalos-Salguero, K. Iltanen, J. Laurikkala, K. Penttinen, J. Hyttinen, K. Aalto-Setälä, and M. Juhola, "Classification of iPSC colony images using hierarchical strategies with support vector machines," in *Proceedings of the 2014 IEEE Symposium on Computational Intelligence and Data Mining*, pp. 86-92, 2014.

[12] M. Juhola, H. Joutsijoki, K. Penttinen, and K. Aalto-Setälä, "Detection of genetic cardiac diseases by Ca^{2+} transient profiles using machine learning methods," *Scientific Reports*, vol. 8, no. 1, 9335, 2018.

[13] M. Juhola, K. Penttinen, H. Joutsijoki, and K. Aalto-Setälä, "Analysis of drug effects on iPSC cardiomyocytes with machine learning," *Annals of Biomedical Engineering*, vol. 49, no. 1, pp. 129-138, 2021.

[14] M.S. Kavitha, T. Kurita, and B.-C. Ahn, "Critical texture pattern feature assessment for characterizing colonies of induced pluripotent stem

cells through machine learning techniques," *Computers in Biology and Medicine*, vol. 94, pp. 55-64, 2018.

[15] M.S. Kavitha, T. Kurita, S.-Y. Park, S.-I. Chien, J.-S. Bae, and B.-C. Ahn, "Deep vector-based convolutional neural network approach for automated recognition of colonies of induced pluripotent stem cells," *PLOS One*, vol. 12, no. 12, Article e0189974, 2017.

[16] C. Kwak and A. Clayton-Matthews, "Multinomial logistic regression," *Nursing Research*, vol. 51, no. 6, pp. 404-410, 2002.

[17] A. L. Lahti, V. J. Kujala, H. Chapman, A.-P. Koivisto, M. Pekkanen-Mattila, E. Kerkelä, J. Hyttinen, K. Kontula, H. Swan, B.R. Conklin, S. Yamanaka, O. Silvennoinen, and K. Aalto-Setälä, "Model for long QT syndrome type 2 using human iPS cells demonstrates arrhythmogenic characteristics in cell culture," *DMM Disease Models and Mechanisms*, vol. 5, no. 2, pp. 220-230, 2012.

[18] C.K. Lam and J.C. Wu, "Clinical trial in a dish: Using patient-derived induced pluripotent stem cells to identify risks of drug-induced cardiotoxicity," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 41, pp. 1019-1031, 2021.

[19] W.-Y. Loh, "Classification and regression trees," *WIREs Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 14-23, 2011.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[21] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 2, pp. 1150-1157, 1999.

[22] P. Musa, F. Al Rafi, and M. Lansani, "A review: Contrast-limited adaptive histogram equalization (CLAHE) methods to help the application of face recognition," in *Proceedings of the 2018 Third International Conference on Informatics and Computing (ICIC)*, 6 pages, 2018.

[23] K. Orita, K. Sawada, R. Koyama, and Y. Ikegaya, "Deep learning-based quality control of cultured human-induced pluripotent stem cell-derived cardiomyocytes," *Journal of Pharmacological Sciences*, vol. 140, no. 4, pp. 313-316, 2019.

[24] T. Piotrowski, O. Rippel, A. Elanzew, B. Nießing, S. Stucken, S. Jung, N. König, S. Haupt, L. Stappert, O. Brüstle, R. Schmitt, and S. Jonas, "Deep-learning-based multi-class segmentation for automated, non-invasive routine assessment of human induced stem cell culture status," *Computers in Biology and Medicine*, vol. 129, Article 104172, 2021.

[25] D.F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, no. 1, pp. 109–118, 1990.

[26] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, and S. Yamanaka, "Induction of pluripotent stem cells from adult human fibroblasts by defined factors," *Cell*, vol. 131, no. 5, pp. 861-872, 2007.

[27] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell*, vol. 126, no. 4, pp. 663-676, 2006.

[28] A. Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *International Journal of Applied Pattern Recognition*, vol. 3, no. 2, pp. 145-180, 2016.

[29] A. Tharwat, T. Gaber, A. Ibrahim, and A.E. Hassanien, "Linear discriminant analysis: A detailed tutorial," *AI Communications*, vol. 30, no. 2, pp. 169-190, 2017.

[30] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," http://www.vlfeat.org, 2008.

[31] X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda G.J. McLachlan, A. Ng, B. Liu, P.S. Yu, Z.-H. Zhou, M. Steinbach, D.J. Hand, and D Steinberg, "Top 10 algorithms in data mining", *Knowledge and Information Systems*, vol. 14, no. 1, pp. 137, 2008.

[32] G. Yue, J. Liao, Y. Wang, L. He, T. Wang, G. Zhou, and B. Lei, "Quality evaluation of induced pluripotent stem cell colonies by fusing multi-source features," *Computer Methods and Programs in Biomedicine*, vol. 208, Article 106235, 2021.