



# Effects of Color Space Transformations on Classification Performance of Sperm Morphology

Mecit Yuzkat<sup>1\*</sup>, Hamza Osman Ilhan<sup>1</sup>, Nizamettin Aydin<sup>1</sup>

<sup>1</sup> Yıldız Technical University, Faculty of Electrical and Electronics, Department of Computer Engineering, Istanbul, Turkey,  
(ORCID: 0000-0003-4808-5181), [myuzkat@yildiz.edu.tr](mailto:myuzkat@yildiz.edu.tr)  
(ORCID: 0000-0002-1753-2703), [hoilhan@yildiz.edu.tr](mailto:hoilhan@yildiz.edu.tr)  
(ORCID: 0000-0003-0022-2247), [naydin@yildiz.edu.tr](mailto:naydin@yildiz.edu.tr)

(5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) 2021 – 21-23 October 2021)

(DOI: 10.31590/ejosat.1013341)

**ATIF/REFERENCE:** Yuzkat, M., Ilhan, H. O., & Aydin, N. (2021). Effects of Color Space Transformations on Classification Performance of Sperm Morphology. *European Journal of Science and Technology*, (29), 70-75.

## Abstract

Infertility is defined by the World Health Organization as the inability of a woman to become pregnant even though the couple had sexual intercourse for one year without any protection. Male and/or female factors might be the reasons for infertility. When diagnosing the male factors, sperm specimens are analyzed in a laboratory environment under certain conditions. The morphological abnormality, characteristic motility and concentration of sperm are examined in the analysis called spermiogram. Spermiogram tests can be done manually by doctors, as well as by using computer-assisted sperm analyzing systems. The importance of computer aided analysis is increasing day by day because visual inspection can give different results from person to person and is costly. In this study, the effect of different color spaces as a preprocessing step is investigated to increase the classification performance of a computer based analyzing approach for sperm morphology. Three sperm morphology data sets abbreviated as SMIDS, HuSHeM and SCIAN-Morpho were used in the experimental tests. Data augmentation was applied on the data sets due to the unbalanced distribution of sperm images among the classes and insufficient data. Then, data sets were converted to two well-known color spaces, LAB and HSV to observe the effects of color space in the classification. MobileNetV2 was utilized as the classification model. In order to indicate the effects of color spaces, results were compared with previously published study where no color transform was implemented. The classification of images in LAB and HSV color spaces had better results than RGB images trained under the same conditions. The maximum classification accuracies of 89%, 85% and 68% were obtained for SMIDS, HuSHeM, SCIAN-Morpho data sets by using the color space transform idea, respectively.

**Keywords:** Deep Learning, MobileNetV2, Computer Aided Sperm Analysis, Morphological Analysis, Color Space Transform.

## Renk Uzayı Dönüşümlerinin Sperm Morfolojisinin Sınıflandırma Performansına Etkisi

### Öz

Dünya Sağlık Örgütü'ne göre kısırlık; çiftlerin herhangi bir koruma olmaksızın bir yıl boyunca cinsel ilişkiye girmelerine rağmen gebeliğin oluşmama durumu olarak tanımlanır. Kısırlığın nedeni erkek ve/veya kadın faktörleri olabilir. Erkek faktörlerin teşhisinde, laboratuvar ortamında belirli koşullar altında sperm hücrelerinin analizi yapılır. Spermiyogram adı verilen analizde sperm morfolojik anormallığı, karakteristik motilitesi ve konsantrasyonu incelenir. Spermiogram testleri doktorlar tarafından manuel olarak yapılabileceği gibi bilgisayar destekli sperm analiz sistemleri kullanılarak da yapılabilmektedir. Görsel incelemenin kişiden kişiye farklı sonuçlar vermesi ve maliyetli olması nedeniyle bilgisayar destekli analizlerin önemi her geçen gün artmaktadır. Bu çalışmada, sperm morfolojisi için bilgisayar tabanlı bir analiz yaklaşımının sınıflandırma performansını artırmak için bir ön işleme adımı olarak farklı renk uzaylarının etkisi araştırılmıştır. Deneysel testlerde SMIDS, HuSHeM ve SCIAN-Morpho olarak kısaltılan üç sperm morfolojisi veri seti kullanılmıştır. Sperm görüntülerinin sınıflar arasındaki dengesiz dağılımı ve yetersiz veri nedeniyle veri setleri üzerinde veri artırma işlemi uygulanmıştır. Daha sonra, renk uzayının sınıflandırmadaki etkilerini gözlemlemek için veri setleri çok iyi bilinen iki renk uzayı olan LAB ve HSV formatlarına dönüştürülmüştür. Sınıflandırma modeli olarak MobileNetV2 kullanılmıştır. Renk uzaylarının etkilerini göstermek için sonuçlar, renk dönüşümünün uygulanmadığı daha önce yayınlanmış çalışma ile karşılaştırılmıştır. LAB ve HSV renk uzaylarında görüntülerin sınıflandırılması, aynı koşullar altında eğitilmiş RGB görüntülerinden daha iyi sonuçlar vermiştir. Renk uzayı dönüşümleri kullanılarak SMIDS, HuSHeM, SCIAN-Morpho veri setleri için sırasıyla %89, %85 ve %68 maksimum sınıflandırma doğruluğu elde edilmiştir.

**Anahtar Kelimeler:** Derin Öğrenme, MobileNetV2, Bilgisayar Destekli Sperm Analizi, Morfolojik Analiz, Renk Uzayı Dönüşümleri.

\* Corresponding Author: Yıldız Technical University, Faculty of Electrical and Electronics, Department of Computer Engineering, Istanbul, Turkey, ORCID: 0000-0003-4808-5181, [myuzkat@yildiz.edu.tr](mailto:myuzkat@yildiz.edu.tr)

## 1. Introduction

Today, the rate of infertile couples is increasing due to environmental, genetic and many other unknown factors. The ratio of only male related factors is 30% of all infertility cases, while 40% of infertility problems have occurred because of male and female based factors together (Agarwal et al. 2015; Pillai and McEleny, 2021). Therefore, the diagnosis of male based problems has a critical impact on the diagnosis of infertility. Sperm specimens are taken and examined under the microscopes in the laboratories for the determination of the male factors. This test is called spermiogram that can be performed by Visual Assessment (VA) technique or using Computer-Assisted Sperm Analysis (CASA) systems. The results of analysis in VA depend on the experience of the assessor/observer, but CASA systems report more objective results because of utilizing the computer based algorithms (Rijsselaere et al., 2012). Additionally, CASA systems are more advantageous than visual evaluation due to utilizing the high image processing techniques reporting the result with more sensitivity, repeatability and practicality (Shi, 2011; Gallardo Bolaños, 2012). Three important features of the sperm as the morphology, motility and concentration are the key parameters to focus in the spermiogram tests. Among these features, the morphological abnormality of the sperm is the best indicator for male factor based infertility problems (Van der Merwe et al., 2005).

There are many studies that have been carried out in the literature on the analysis of sperm morphology in recent years. One of these studies is morphological classification of sperm images using a dictionary learning method and Support Vector Machine (SVM). The authors used two differently arranged data sets namely HuSHeM and SCIAN-MorphoSpermGS in their studies. Classification accuracies were measured as 92% and 62% accuracies for HuSHeM and SCIAN-MorphoSpermGS datasets by using the SVM classifier based dictionary learning approach, respectively. However, automatic classification was not performed in this study. Instead of using the original images in the HuSHeM data set, they manipulated images by applying pre-processes such as cropping and changing the direction (Shaker et al., 2017).

In another study, the researchers established a gold standard for classification of the SCIAN-MorphoSpermGS data set. They used three different shape-based descriptors (Hu moments, Zernike moments, Fourier) as well as four supervised learning methods (K-Nearest Neighbors, Naive Bayes, Decision Trees and SVM) to create the standard. None of these approaches appeared to overcome the sperm head classification problem. The optimal result was obtained as 49% using the Fourier descriptor and SVM (Chang et al., 2017).

Riordon et al. obtained 94% and 62% accuracies for HuSHeM and SCIAN-MorphoSpermGS data sets, respectively, by modifying the VGG16 network. In their work, similar to Shaker (Shaker et al., 2017), they manually manipulated images such as cropping and rotating before classification. Therefore, the proposed classification approach is not fully automatic. (Riordon et al., 2019).

Yüzkat et al. applied various interpolation and data augmentation methods to increase the classification performance of the SCIAN-MorphoSpermGS data set. Then, 62% classification accuracy was achieved by using a modified VGG-

19 network (Yüzkat et al., 2020). In a more recent study, they presented a fully automated analyzing approach for classifying sperm images. In their study, they created six CNN models and combined the results of the models with fusion methods (hard and soft). They obtained 90.73%, 85.18% and 71.91% accuracy for the SMIDS, HuSHeM and SCIAN-MorphoSpermGS datasets, respectively (Yüzkat et al., 2021).

In another fully automatic classification approach for sperm images, Ilhan et al. initially implemented a special preprocessing approach to SMIDS and HuSHeM data sets. They used an automated preprocessing technique called as directional masking which had been published in (Ilhan et al., 2019). After applying the directional masking technique on images, they extract the SURF and MSER descriptive features of the preprocessed images. Then, extracted features are classified by SVM. They increased their previous success rate for SMIDS from 84% to 86%. For HuSHeM, they achieved an accuracy of 85% (Ilhan et al., 2020a).

Ilhan et al. used MobileNetV2 in another study similar to this study. But instead of using MobileNetV2 as a pretrained network, they retrained the network from the stretch by using SMIDS. By applying different data augmentation techniques, they have reached 87% accuracy. The disadvantage of the work is that it takes a lot of time to train such a large network from the stretch (Ilhan et al., 2020b).

In order to investigate the mobile platform based pre-trained network performances for the classification of sperm images, Tortumlu and Ilhan, compared the classification performances of MobileNet V1 and V2 over three sperm morphology data sets (Tortumlu and Ilhan, 2020). They achieved 77%, 88% and 67% classification accuracies for HuSHeM, SMIDS and SCIAN-Morpho GS datasets, respectively.

In this study, we aimed to increase the classification performance of the mobile platform based network, MobileNetV2. In the previously published study (Tortumlu and Ilhan, 2020), classification was made using RGB formatted sperm images. In the presented study, we investigated the effect of different color spaces in which the morphology of sperm images appear more distinctive in the classification. We converted the RGB format images used in the published study into two different image spaces as LAB and HSV. We classified these images using the MobileNetV2 network and compared them with the previous work.

## 2. Material and Method

In the previously published study (Tortumlu and Ilhan, 2020), mobile platform-based networks, MobileNetV1 and MobileNetV2, were applied to three original formats of data sets. MobileNetV2 was defined as the most successful network according to the obtained results. In this regard, MobileNetV2 is utilized in this study as a classifier to investigate the effect of color space conversion preprocess over the classification performance of sperm images. The fine-tuning procedure of MobileNetV2 has been performed to LAB and HSV formats instead of RGB formats of the three data sets. Within the scope of this study; epoch numbers, data augmentation sizes and other parameters were selected the same way as in the previously published study. The general flowchart of the presented study is demonstrated in Figure 1.

Our study was carried out with MATLAB 2020a (MATLAB, 2020) on the Windows Operating System (Version 10). NVIDIA Geforce GTX 970 was used to speed up training

times. In addition to the GPU module, 8 GB of local ram and Intel i5 CPU with 3.2 GHz processing speed were also used on the test platform.

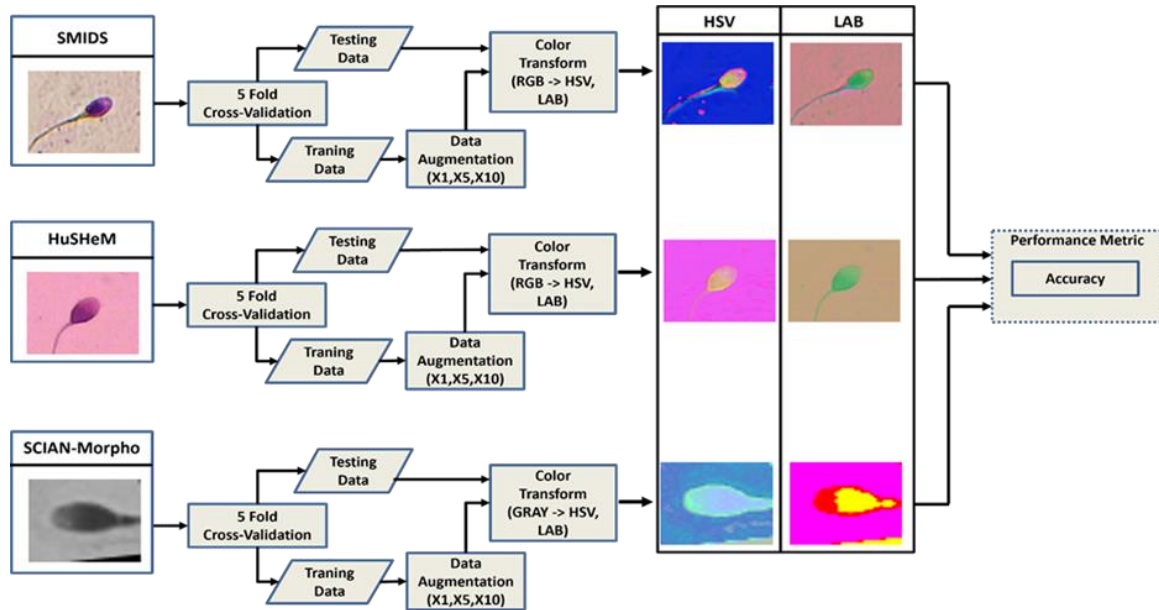


Figure 1. Flowchart of the study

## 2.1. Data Set Information

Three data sets, SMIDS (Sperm Morphology Image Data Set), HuSHeM (Human Sperm Head Morphology) and SCIAN-MorphoSpermGS, were used in this study. SMIDS and HuSHeM datasets are obtained in RGB color space. On the other hand, the

images in the SCIAN-MorphoSpermGS dataset are in grayscale monochromatic format. The image distributions per class of the data sets are given in Figure 2.

SMIDS			HuSHeM			SCIAN-Morpho		
Labels	# of images	Example	Labels	# of images	Example	Labels	# of images	Example
Normal	1021		Normal	54		Normal	100	
Abnormal	1005		Tapered	53		Tapered	228	
Non-Sperm	974		Pyriform	57		Pyriform	76	
			Amorphous	52		Small	72	
						Amorphous	656	
<b>Total</b>	<b>3000</b>		<b>Total</b>	<b>216</b>		<b>Total</b>	<b>1132</b>	

Figure 2. The image and class distribution of data sets

The images in the SMIDS are equally distributed into three classes as Normal, Abnormal, and Non-Sperm. The HuSHeM data set has four classes: Normal, Tapered, Pyriform and Amorphous. Although the image distribution between classes is balanced, the total number of images is insufficient for a deep learning based classification approach. Lastly, The SCIAN-MorphoSpermGS data set has 5 classes: Normal, Tapered, Pyriform, Small and Amorphous. The distribution of images over classes is not balanced in this data set. For example, 656 images

are labelled as the amorphous class, while the small class has only 72 images. In order to create the training and test sets, K-Fold Cross validation schema was used where the K is assigned as 5. In this regard, 2400, 171, 906 images are located in the training set of SMIDS, HuSHeM, and SCIAN-Morpho data sets, respectively. Rest of the images (600, 45 and 226) are reserved for the test set. According to the implementation of K-Fold Cross Validation technique, images are rotated to create new sets for each fold. Therefore the test accuracy reflects the classification

performance of the dataset. In addition, data augmentation processes were carried out in the training set of these data sets.

**2.2. Data Augmentation and Epoch Size Analysis**

In the classification problems by deep learning techniques, having enough data allows the training phase to be better (Lee et al., 2017). Using the different variations of the images in the training phase of the model eliminates the memorizing effect in the classification of the test set. These variations can be generated by using the data augmentation techniques such as resizing, scrolling, scaling, cropping etc. While augmenting the data, the parameters of the augmentation techniques are carefully selected in order not to deform the original sperm images and mislead the classification output. For example, with improperly applied data augmentation techniques, normal sperm images may be perceived as abnormal. As an illustration of the implemented data augmentation types, six data augmentation processes applied on the same images for three datasets are given in Figure 3. These operations are a) height shift + shear, b) horizontal flip + zoom, c) width shift + height shift, d) horizontal flip + height shift, e) rotation, f) height shift. To avoid distortion of images, zooming, cropping, and panning were limited to 10%, 5%, and 15% of the original image sizes for all three data sets, respectively. Besides the original datasets (nominated as X1),

two scales of augmentation have been applied as X5 and X10 which indicate the 5 and 10 times augmented datasets.

Additionally, preliminary tests were performed by applying different epoch and mini batch sizes with different activation functions. As a result of these tests, the parameters giving the best accuracy values were determined. While determining these parameters, we tested the values frequently used in the literature. For example, while determining the minibatch size, we used 4, 8, 16, 32 mini batch sizes in our previous study (Yüzkat et al., 2021), and we generally achieved the best result by using 8 for each data set. In this study, we set the minibatch parameter to 8 according to our preliminary tests. In terms of activation function, Adam was used in the experimental tests similar to the previous study (Tortumlu and Ilhan, 2020).

We also researched the positive effect of the different epoch sizes used in the training phase on the classification. High epoch sizes (training phase iteration steps) can cause greater processing time requirements as well as an overfitting problem. Low epoch sizes may cause the network not to train fully. For this reason, training was carried out with epoch sizes of 10, 30 and 50 to observe the effect of epoch sizes on the classification performances.

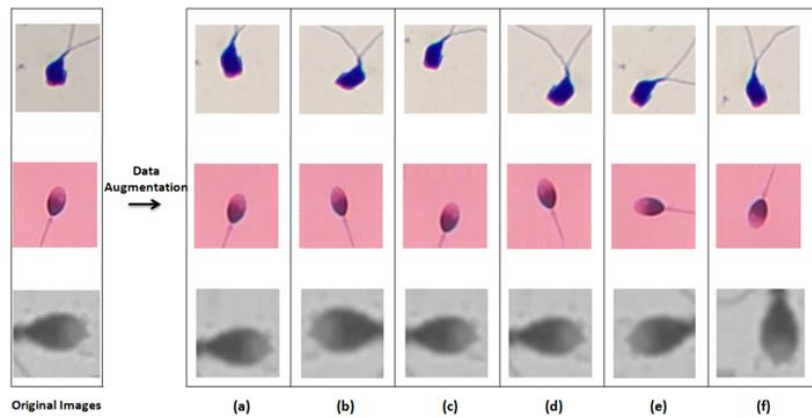


Figure 3. Data augmentation techniques applied for each data set

**2.3. Mobile Platform Based Convolutional Neural Network**

In this study, the classification performance of one of the mobile platform based convolutional neural network models, MobileNetV2, was evaluated over three sperm morphology data sets. MobileNetV2 is a lightweight network and used in terms of transfer learning idea in this study. We modified the classification layer according to the class structure of the utilized data sets, using a basic convolution layer and a Softmax classifier layer. Then, we applied the model to different sized data sets using the Adam optimizer. The main reason for using

MobileNetV2 is that it is more suitable for non-computer mobile applications and gives better results than other models such as MobileNetV1, InceptionV3, DenseNet121 (Sandler et al., 2018; Xiang et al., 2019). The main criterion in the evaluation will be to ensure the highest accuracy and minimum time complexity in classification. LAB and HSV color space conversion have been performed as a preprocessing step in order to observe the effects of different color spaces in the classification performances. The training parameters of the MobileNetV2 network are set according to Table 1.

Table 1. Parameters used in the study

Parameter Name	Configuration
Batch Size	8
Optimizer	Adamax
Learning Rate	0.0001
Augmentation Sizes	X1 (Original Data), X5, X10
Epoch Sizes	10, 30, 50



### 2.4. Color Space Transform

In the field of image processing, especially in image segmentation, the use of different color spaces has produced results with different accuracy. Therefore, many researches have been performed about the selection of the appropriate color space in image processing. Generally, LAB and HSV are the two most important color spaces chosen (Bora et al., 2015). Made a performance comparison between these two color spaces for image segmentation. They used parameters such as Mean-Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) to measure the accuracy of their performance. In their study, they stated that the HSV color space outperformed LAB. In our study, RGB-formatted SMIDS, HuSHeM data sets were converted to LAB and HSV formats. In addition, the SCIAN-Morpho data set in grayscale format was converted to LAB and HSV. Then, these data sets were classified by using MobileNetV2 under the same conditions as in the previously published study (Tortumlu and Ilhan, 2020).

### 3. Results and Discussion

The classification performance of MobileNetV2 was measured over different data augmentation scales of original

data sets. Additionally, several epoch sizes were utilized in the training phase of the network to achieve the maximum accuracy. In the previously published study, while all three data sets were given as input to mobile networks, 3-channel RGB color space format was used. In our study, the classification results of HSV and LAB color space converted data sets were compared with the results of the RGB color space used in the previously published study (Tortumlu and Ilhan, 2020). The classification performance results are given in Table 2 and 3 for the images converted from RGB to HSV and LAB, respectively.

According to Table 2 and Table 3, the classification performances obtained for the augmented data sets (X5 and X10) are higher than the original (X1) data sets in terms of both HSV and LAB conversion. The highest classification success after HSV conversion was obtained as 81%, 88% and 68% accuracy rates for HuSHeM, SMIDS and SCIANMorphoSpermGS datasets, respectively. In the classification of LAB converted images, maximum performances as 85%, 89% and 65% accuracy rates were obtained for HuSHeM, SMIDS and SCIANMorphoSpermGS datasets.

Table 2. Comparison of accuracy values on data sets of two studies ( RGB->HSV )

Aug/Epo	MobileNetV2 (Tortumlu and Ilhan, 2020) (RGB Format)									MobileNetV2 (Results of Our Study) (HSV Format)								
	HuSHeM			SMIDS			SCIAN			HuSHeM			SMIDS			SCIAN		
	10	30	50	10	30	50	10	30	50	10	30	50	10	30	50	10	30	50
X1	38	44	39	84	85	84	48	54	59	67	74	72	85	86	85	65	64	65
X5	62	73	75	85	<b>87</b>	87	54	62	62	78	80	71	88	87	88	67	<b>68</b>	67
X10	72	<b>77</b>	77	86	86	87	65	<b>66</b>	66	81	<b>81</b>	78	88	<b>88</b>	87	66	65	66

Table 3. Comparison of accuracy values on data sets of two studies ( RGB->LAB )

Aug/Epo	MobileNetV2 (Tortumlu and Ilhan, 2020) (RGB Format)									MobileNetV2 (Results of Our Study) (LAB Format)								
	HuSHeM			SMIDS			SCIAN			HuSHeM			SMIDS			SCIAN		
	10	30	50	10	30	50	10	30	50	10	30	50	10	30	50	10	30	50
X1	38	44	39	84	85	84	48	54	59	76	76	75	87	88	86	60	62	63
X5	62	73	75	85	<b>87</b>	87	54	62	62	80	<b>85</b>	85	89	<b>89</b>	88	65	64	62
X10	72	<b>77</b>	77	86	86	87	65	<b>66</b>	66	82	84	83	89	89	88	64	<b>65</b>	62

When both tables are examined, the best results are obtained when 30 epochs were used in the training. In other words, faster and higher results can be obtained by using 30 epochs without waiting more time for 50 epochs. Additionally, models seem overfitting when training is set to 50 epochs especially for HSV converted HuSHeM and LAB converted SCIANMorphoSpermGS datasets.

When comparing the classification performance of converted datasets to original forms (RGB), higher accuracy rates were achieved in both HSV and LAB color spaces than RGB data sets. The highest accuracy values for HuSHeM and SMIDS were obtained as 85% and 89% after the LAB color space conversion preprocess, respectively. For the SCIAN data

set, the maximum accuracy was measured as 68% in the classification of HSV converted images. Thus, HuSHeM and SMIDS are better classified in the LAB color space, while SCIAN is better classified in the HSV color space.

### 4. Conclusions and Recommendations

The number of male-based infertility is increasing day by day. Today, CASA systems are very popular in sperm analysis because of eliminating the variability problem between the observers in traditional visual analysis. However, many laboratories can not afford those systems due to high costs. Additionally, these systems do not provide a fully automated analysis. Therefore, instead of using compact CASA systems,

different image processing based studies have been performed to create more reasonable solutions. In the presented study, three publicly available sperm morphology data sets abbreviated as SMIDS, HuSHeM and SCIAN-Morpho were aimed to classify with the high accuracy rates. These data sets were augmented by applying different data augmentation processes due to the lack of data. Then, the SMIDS and HuSHeM datasets were converted from RGB color spaces to LAB and HSV. The SCIAN-MorphoSpermGS data set was converted from grayscale. The classification performance of one of mobile platform based convolutional neural networks, MobileNetV2, has been measured over the color converted data sets in order to observe the color conversion effect on the classification. Different epoch and augmented sizes were also tested. As a result, using MobileNetV2 in the classification of images with HSV and LAB color space resulted in higher accuracies than RGB color space images. The obtained accuracy differences are 8%, 2%, and 2% for HuSHeM, SMIDS, and SCIAN-Morpho data sets, respectively.

The disadvantage of our study was that we could not achieve sufficient success in the SCIAN-Morpho data set. In our next work, we will use different color spaces in addition to HSV and LAB. We intend to obtain new results by training the data sets to be obtained separately with different deep learning networks. It will be to use the multiple results to be obtained with fusion methods for more stable classification. Thus, in our study, we aim to achieve sufficient classification success in the SCIAN-Morpho data set and to increase the success rates obtained in the SMIDS and HuSHeM data sets.

## References

- Agarwal, A., Mulgund, A., Hamada, A., & Chyatte, M. R. (2015). A unique view on male infertility around the globe. *Reproductive biology and endocrinology*, 13(1), 1-9.
- Bora, D. J., Gupta, A. K., & Khan, F. A. (2015). Comparing the performance of L\* A\* B\* and HSV color spaces with respect to color image segmentation. *arXiv preprint arXiv:1506.01472*.
- Chang, V., Garcia, A., Hirschfeld, N., & Härtel, S. (2017). Gold-standard for computer-assisted morphological sperm analysis. *Computers in biology and medicine*, 83, 143-150.
- Gallardo Bolaños, J. M., Miró Morán, Á., Balao da Silva, C. M., Morillo Rodríguez, A., Plaza Dávila, M., Aparicio, I. M., ... & Peña, F. J. (2012). Autophagy and apoptosis have a role in the survival or death of stallion spermatozoa during conservation in refrigeration. *PloS one*, 7(1), e30688.
- Ilhan, H. O., Serbes, G., & Aydin, N. (2019). Automatic directional masking technique for better sperm morphology segmentation and classification analysis. *Electronics Letters*, 55(5), 256-258.
- Ilhan, H. O., Serbes, G., & Aydin, N. (2020a). Automated sperm morphology analysis approach using a directional masking technique. *Computers in Biology and Medicine*, 122, 103845.
- Ilhan, H. O., Sigirci, I. O., Serbes, G., & Aydin, N. (2020b). A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods. *Medical & biological engineering & computing*, 58(5), 1047-1068.
- Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. *Korean journal of radiology*, 18(4), 570-584.
- MATLAB Version 9.8.0.1323502 (R2020a), The Mathworks, Inc., Natick, Massachusetts (2020)
- Pillai, R. N., & McEleny, K. (2021). Management of male infertility. *Obstetrics, Gynaecology & Reproductive Medicine*.
- Rijsselaere, T., Van Soom, A., Maes, D., & Nizanski, W. (2012). Computer-assisted sperm analysis in dogs and cats: An update after 20 years. *Reproduction in Domestic Animals*, 47, 204-207.
- Riordon, J., McCallum, C., & Sinton, D. (2019). Deep learning for the classification of human sperm. *Computers in biology and medicine*, 111, 103342.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
- Shaker, F., Monadjemi, S. A., Alirezaie, J., & Naghsh-Nilchi, A. R. (2017). A dictionary learning approach for human sperm heads classification. *Computers in biology and medicine*, 91, 181-190.
- Shi, X. D., Bi, H. J., Fu, H. L., Li, L. Y., Liu, D. K., & Li, J. M. (2011). Effect of low-dose fenvalerate on semen quality capacitation in adult mice. *Chinese medical journal*, 124(10), 1529-1533.
- Tortumlu, O. L., & Ilhan, H. O. (2020, November). The Analysis of Mobile Platform based CNN Networks in the Classification of Sperm Morphology. In *2020 Medical Technologies Congress (TIPTEKNO)* (pp. 1-4). IEEE.
- Van der Merwe, F. H., Kruger, T. F., Oehninger, S. C., & Lombard, C. J. (2005). The use of semen parameters to identify the subfertile male in the general population. *Gynecologic and obstetric investigation*, 59(2), 86-91.
- Xiang, Q., Wang, X., Li, R., Zhang, G., Lai, J., & Hu, Q. (2019, October). Fruit image classification based on Mobilenetv2 with transfer learning technique. In *Proceedings of the 3rd International Conference on Computer Science and Application Engineering* (pp. 1-7).
- Yüzkat, M., Ilhan, H. O., & Aydin, N. (2020, November). Morphological Classification of Low Quality Sperm Images Using Deep Learning Networks. In *2020 Medical Technologies Congress (TIPTEKNO)* (pp. 1-4). IEEE.
- Yüzkat, M., Ilhan, H. O., & Aydin, N. (2021). Multi-Model CNN Fusion for Sperm Morphology Analysis. *Computers in Biology and Medicine*, 104790. <https://doi.org/10.1016/j.combiomed.2021.104790>