

# Melanoma Recognition

Michal Haindl<sup>a</sup> and Pavel Žid<sup>b</sup>

*The Institute of Information Theory and Automation, Czech Academy of Sciences,  
Pod Vodárenskou věží 4, Prague, Czech Republic*

**Keywords:** Skin Cancer Recognition, Melanoma Detection, Circular Markov Random Field Model.

**Abstract:** Early and reliable melanoma detection is one of today's significant challenges for dermatologists to allow successful cancer treatment. This paper introduces multispectral rotationally invariant textural features of the Markovian type applied to effective skin cancerous lesions classification. Presented texture features are inferred from the descriptive multispectral circular wide-sense Markov model. Unlike the alternative texture-based recognition methods, mainly using different discriminative textural descriptions, our textural representation is fully descriptive multispectral and rotationally invariant. The presented method achieves high accuracy for skin lesion categorization. We tested our classifier on the open-source dermoscopic ISIC database, containing 23 901 benign or malignant lesions images, where the classifier outperformed several deep neural network alternatives while using smaller training data.


## 1 INTRODUCTION


Diagnosing skin diseases is complicated. At least 3,000 identified varieties of skin diseases (Kawahara and Hamarneh, 2019) with a prevalence that varies by condition. Skin diagnosis is usually determined using a biopsy, which, however lengthy, costly, uncomfortable, and may introduce potential infectious complications to the patient. Image-based automatic skin diagnosis potentially avoids all these difficulties. Automatic skin cancer recognition is a challenging but proper application that can help dermatologists in early cancer detection. Melanoma accounted for 41% skin-related deaths in the USA in 2013 (Kawahara and Hamarneh, 2019) where every hour one person dies from melanoma, while the highest melanoma rate is in Australia (34 000 skin cancers every year) and New Zealand (Zhang et al., 2020). Thus the early melanoma diagnosis is of the utmost importance. Diagnosis of skin lesion type is a challenging task even for a skilled dermatologist. The use of dermoscopy imaging devices significantly improved the quality of early melanoma detection. Further improvement is possible by the use of computer-assisted diagnosis.

Automatic skin lesion categorization allows identification or learning of skin lesion types possible without specific medical knowledge. Recent stud-

ies show that recognition systems can match or even outperform clinicians in the diagnosis of individual skin lesion images in controlled reader studies (Kawahara and Hamarneh, 2019; Rotemberg et al., 2021). Results comparison is difficult due to different data sets or their subset used in different studies and often not appropriately described. (Ballerini et al., 2013) reached 74% accuracy for five classes using the hierarchical k-NN classifier using mean and covariance color matrices, and 12 features derived from gray-level co-occurrence matrices (3888 features) on their dataset with 960 images acquired from Canon EOS 350D SLP camera. Authors (Gomez and Herrera, 2017) used an SVM classifier with histogram and gray-level co-occurrence matrices-based features and reached 70% accuracy on UDA, MSK, and SONIC data sets. The gray level co-occurrence features are similarly used as texture features in (Shakourian Ghalejoogh et al., 2019).

Several algorithms (Akram et al., 2020; Ballerini et al., 2013; Esteva et al., 2017; Favole et al., 2020; Gessert et al., 2020; Hosny et al., 2020; Kawahara et al., 2016; Khan et al., 2021; Mahbod et al., 2019; Nida et al., 2019; Tschandl et al., 2019; Zhang et al., 2019; Zhang et al., 2020) are using convolutional neural networks (CNN). (Kawahara et al., 2016) reached 85% accuracy for five classes on AlexNet (4096 features) on the Dermofit (Ballerini et al., 2013) 1300 images. (Esteva et al., 2017) used GoogleNet Inception

<sup>a</sup>  <https://orcid.org/0000-0001-8159-3685>

<sup>b</sup>  <https://orcid.org/0000-0001-8249-1701>

v3 CNN and achieved 72% accuracy for three classes and 55% accuracy for nine classes on 129450 images. (Tschandl et al., 2019) use CNN for lesion segmentation from the HAM10000 dataset (Tschandl et al., 2018). (Zhang et al., 2019) used an attention residual learning convolutional neural network evaluated on the ISIC-skin 2017 dataset (Codella et al., 2018) with 87% accuracy for the melanoma classification.

We use the (ISI, ) skin image database chosen by Favole et al. (Favole et al., 2020). The authors (Favole et al., 2020) examine the performance of AlexNet (Krizhevsky et al., 2012), Inception-V1 (a.k.a. GoogLeNet) (Szegedy et al., 2015) and Resnet50 (He et al., 2016) CNNs for the classification problem of skin lesions.

Our study tests the circular 2D causal auto-regressive adaptive random (2DSCAR) field model and compares the results with results published in (Favole et al., 2020). We use 23901 dermoscopic images of lesions from the ISIC image database. We have chosen the 2018 JID Editorial, HAM10000, MSK, SONIC, and UDA datasets. The presented contribution is the accuracy improvement of skin melanoma recognition while using a smaller training set and faster learning than the alternative deep neural net approach.

## 2 CIRCULAR MARKOVIAN TEXTURE REPRESENTATION

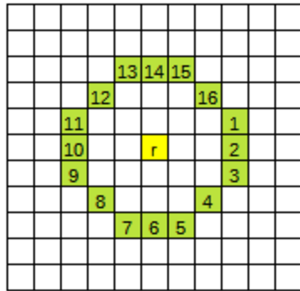


Figure 1: The octagonal, circular path. The numbers mark the site order in which the pixels  $s$ , i.e.,  $I_s^{cs}$  contextual neighborhoods are traversed. Furthermore, the center pixel  $r$ , for which the statistics are computed, is marked by the yellow square.

The circular 2D causal auto-regressive adaptive random (2DSCAR) field model (Fig. 1) is a generalization of the directional 2DCAR model (Haindl, 2012) to the rotationally invariant form, which was introduced in (Remeš and Haindl, 2018) for bark classification application. The model’s contextual neighbor index shift set denoted  $I_r^{cs}$  is functional. The model

for  $d$  spectral bands can be defined in the following matrix form:

$$Y_r = \gamma Z_r + e_r, \quad (1)$$

where  $\gamma = [A_1, \dots, A_\eta]$  is the parameter matrix,  $A_i = \text{diag}[a_{i1}, \dots, a_{id}] \forall i$ ,  $a_{ij}$  are unknown parameters to be estimated (2),  $\eta = \text{cardinality}(I_r^{cs})$ ,  $r = [r_1, r_2]$  is spatial multi-index ( $r_1$  row and  $r_2$  column indices) denoting history of movements on the lattice  $I$ ,  $e_r$  denotes the driving white Gaussian noise vector with zero mean and a constant but unknown covariance matrix  $\Sigma$ , and  $Z_r$  is a neighborhood support vector of multispectral pixels  $Y_{r-s}$  where  $s \in I_r^{cs}$ .

All 2DSCAR model statistics can be efficiently analytically estimated as proven in (Haindl, 2012). The Bayesian parameter estimation (conditional mean value)  $\hat{\gamma}$  can be accomplished using fast, numerically robust and recursive statistics (Haindl, 2012), given the known 2DSCAR process history

$$Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1, Z_t, Z_{t-1}, \dots, Z_1\} :$$

$$\hat{\gamma}_{t-1}^T = V_{zz(t-1)}^{-1} V_{zy(t-1)}, \quad (2)$$

$$V_{t-1} = \tilde{V}_{t-1} + V_0, \quad (3)$$

$$\begin{aligned} \tilde{V}_{t-1} &= \begin{pmatrix} \sum_{u=1}^{t-1} Y_u Y_u^T & \sum_{u=1}^{t-1} Y_u Z_u^T \\ \sum_{u=1}^{t-1} Z_u Y_u^T & \sum_{u=1}^{t-1} Z_u Z_u^T \end{pmatrix} \\ &= \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}_{zy(t-1)}^T \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{zz(t-1)} \end{pmatrix}, \quad (4) \end{aligned}$$

where  $V_{t-1}$  is the data gathering matrix,  $V_0$  is a positive definite initialization matrix (see (Haindl, 2012)). We introduce a new octagonal traversing order multi-index  $t$  of the sequence of multi-indices  $r$ , to simplify notation, which depends on the selected model movement in the underlying lattice  $I$  (e.g.,  $\mathcal{T} = \{t_1, t_1 + (1;0), t_1 + (2;0), \dots, t_{16} + (-1; -1)\}$  for Fig. 1). The optimal functional causal contextual neighbourhood  $I_r^{cs}$  (Fig. 2) can be solved analytically by a straightforward generalization of the Bayesian estimate derived in (Haindl, 2012). We did not optimize the neighbourhood  $I_r^{cs}$  but used its fixed form Fig. 2 to simplify and speed up our experiments. However, if this neighborhood is optimized, we can expect further accuracy improvement. The model can be easily applied also to various synthesis and restoration applications. The 2DSCAR model pixel-wise synthesis is a direct application of the equation (1) fed from a Gaussian noise generator for any 2DSCAR model.

### 2.1 Circular Models

The 2DSCAR model moves ( $r$ ) on the circular path on the lattice  $I$  as is illustrated in Fig. 1. The causal

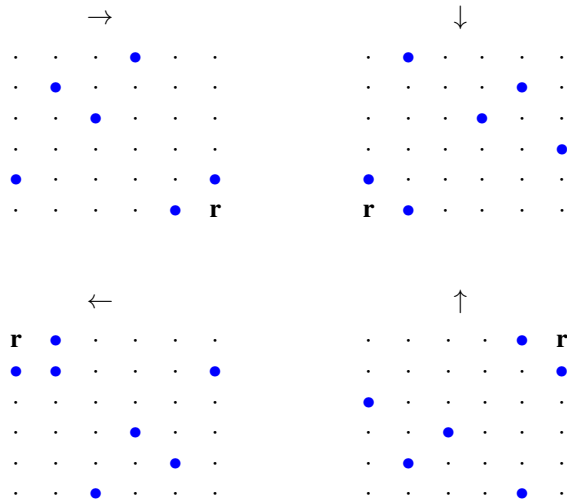


Figure 2: The applied fixed causal functional contextual neighbourhood  $I_r^{cs}$  in four selected directions,  $\bullet$  are the contextual neighbours,  $r$  is the neighbourhood location index. Upper left: rightwards (Fig.1-13,14,15), upper right: downwards (Fig.1-1,2,3), bottom left leftwards (Fig.1-5,6,7), bottom right upwards (Fig.1-9,10,11), respectively.

neighborhood  $I_r^c$  has to be transformed to be consistent for each direction in the traversed path, as denoted in Fig. 2. The paths used can be arbitrary as long as they keep transforming the causal neighborhood into  $I_r^{cs}$  in such a way that the model has visited all neighbors of a control pixel  $r$ . Thus these neighbors are known from the previous steps. We shall call all these causal paths as circulars further on. In this paper, we present the octagonal type of path - (Fig. 1). However, alternatively, a circular path can be used as well. The parameters for the center pixel (the yellow square in Fig. 1) of the circular are estimated after the whole path is completed. Since this model's equations do not need the whole history of movement through the image but only the local neighborhood of a single circular, the 2DSCAR models can be easily parallelized. This memory restriction is advantageous in comparison to the standard directional CAR models (Haindl, 2012). The 2DSCAR models exhibit rotational invariant properties for the circular shape paths, thanks to the CAR model's memory of all the visited pixels. Additional prior contextual information can be easily incorporated if every initialization matrix  $V_0 = V_{t-1}$ , for example, this matrix can be initialized from the previous data gathering matrix.

## 2.2 Multispectral Rotationally Invariant Features

We analyzed the 2DSCAR model around all pixels with the vertical and horizontal stride of 2 to speed

up the computation for feature extraction. The following  $\alpha_1, \alpha_2, \alpha_3$  illumination invariant features initially derived for the 3DCAR model (Haindl, 2012) were adapted for the 2DSCAR model:

$$\alpha_1 = 1 + Z_r^T V_{zz}^{-1} Z_r, \quad (5)$$

$$\alpha_2 = \sqrt{\sum_r (Y_r - \hat{y}Z_r)^T \lambda_r^{-1} (Y_r - \hat{y}Z_r)}, \quad (6)$$

$$\alpha_3 = \sqrt{\sum_r (Y_r - \mu)^T \lambda_r^{-1} (Y_r - \mu)}, \quad (7)$$

where  $\mu$  is the mean value of vector  $Y_r$  and

$$\lambda_{t-1} = V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1}. \quad (8)$$

The inversion data gathering matrix  $V_{zz(t-1)}^{-1}$  is updated in its square-root Cholesky factor to guarantee numerical stability for computed model statistics (Haindl, 2012). Additional used texture features are also the estimated trace of  $\gamma$  parameters, the posterior probability density (Haindl, 2012)

$$p(Y_r | Y^{(r-1)}, \hat{y}_{r-1}) = \quad (9)$$

$$\frac{\Gamma(\frac{\beta(r)-\eta+3}{2})}{\Gamma(\frac{\beta(r)-\eta+2}{2}) \pi^{\frac{1}{2}} |\lambda_{(r-1)}|^{\frac{1}{2}} (1 + X_r^T V_{x(r-1)}^{-1} X_r)^{\frac{1}{2}}} \left( 1 + \frac{(Y_r - \hat{y}_{r-1} X_r)^T \lambda_{(r-1)}^{-1} (Y_r - \hat{y}_{r-1} X_r)}{1 + X_r^T V_{x(r-1)}^{-1} X_r} \right)^{-\frac{\beta(r)-\eta+3}{2}},$$

$\beta(r) = r + \eta - 2$ , and the absolute error of the one-step-ahead model prediction (Haindl, 2012):

$$\begin{aligned} Abs(GE) &= \left| E \left\{ Y_r | Y^{(r-1)} \right\} - Y_r \right| \\ &= |Y_r - \hat{y}_{r-1} X_r|. \end{aligned} \quad (10)$$

Fig. 3 illustrates 15 selected features computed from the HAM database ISIC\_0024313 malignant lesion.

## 3 SKIN LESION CLASSIFIER

The algorithm starts with image subsampling to the width of 512 px (for larger images) while keeping the aspect ratio to speed up the feature extraction part. This subsampling ratio depends on application data; it is a compromise between the algorithm efficiency and its recognition rate. Every pixel has extracted features  $\theta$ , as described in Sec. 2. The resulting feature space indexed on the lattice  $I$  is assumed to be governed by the multivariate Gaussian distribution. The  $n$  estimated Gaussian parameters then represent every training image sample:

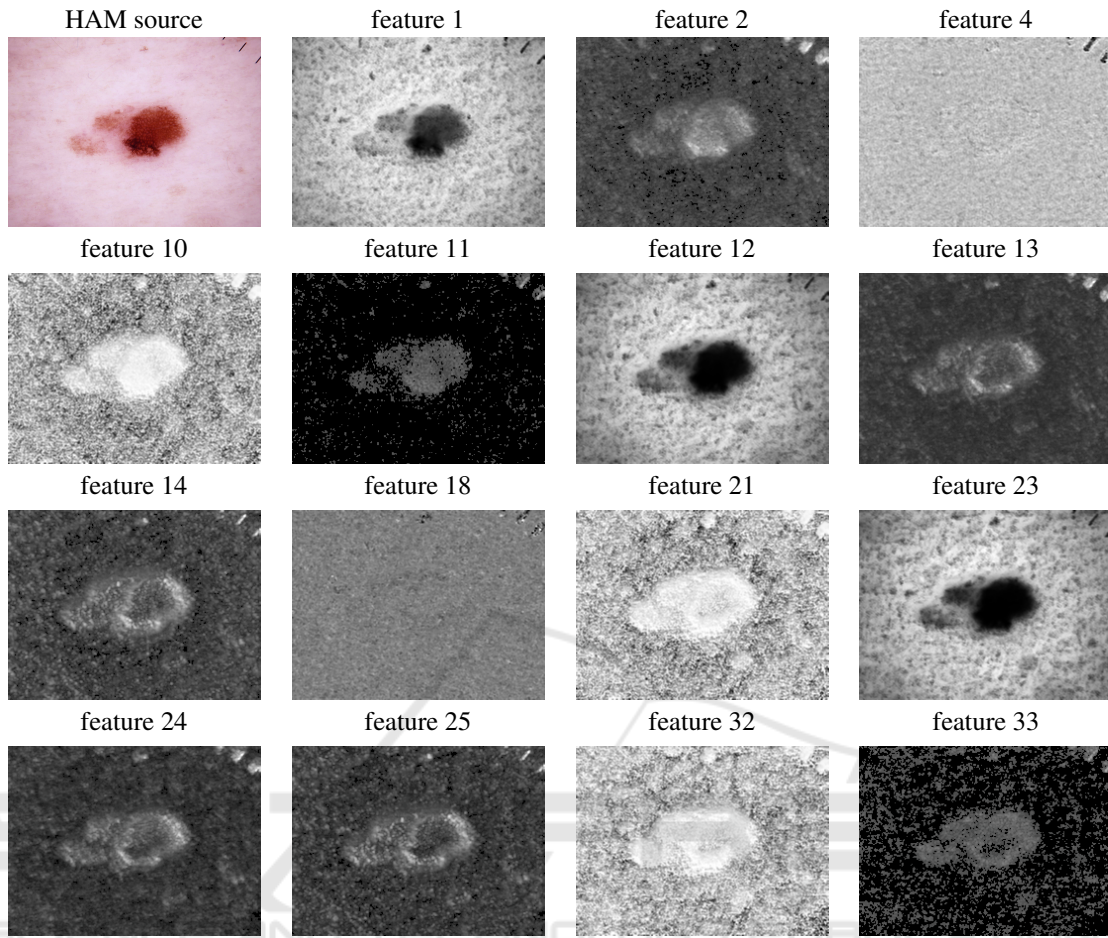


Figure 3: Examples of the selected 2DSCAR features computed from the HAM ISIC\_0024313 malignant texture.

$$\mathcal{N}(\theta|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp \left\{ -\frac{1}{2} (\theta - \mu)^T \Sigma^{-1} (\theta - \mu) \right\} . \quad (11)$$

In the classification step, the Gaussian distribution parameters are estimated for the classified image in the same way. The classified image parameters are then compared with all the distributions from the training samples set using the 1 nearest-neighbor classifier (1-NN) and the Jeffreys divergence (14) as the measure. The KL divergence is a probability distribution non-symmetric similarity measure between two distributions; it is defined as:

$$D(f(x)||g(x)) \stackrel{def}{=} \int f(x) \log \frac{f(x)}{g(x)} dx , \quad (12)$$

where  $f(x), g(x)$  are the compared probability densities.

The KL divergence for the Gaussian distribution

data model can be solved analytically:

$$D(f(x)||g(x)) = \frac{1}{2} \left( \log \frac{|\Sigma_g|}{|\Sigma_f|} + tr(\Sigma_g^{-1} \Sigma_f) - d \right) + \frac{1}{2} (\mu_f - \mu_g)^T \Sigma_g^{-1} (\mu_f - \mu_g) . \quad (13)$$

We use the Jeffreys divergence, which the symmetrized variant of the Kullback-Leibler divergence:

$$D_s(f(x)||g(x)) = \frac{D(f(x)||g(x)) + D(g(x)||f(x))}{2} . \quad (14)$$

The selected class is the class of a training image with the lowest Jeffreys divergence from the tested image. The primary benefit of our method is the significant compression of the training database into the Gaussian distribution parameters (as we extract only about  $n = 40$  features, depending on the chosen neighborhood, we need to store 40 real numbers for the mean and  $40 \times 40$  numbers for the covariance matrix). The subsequent comparison with the training



Figure 4: Examples of skin lesion images (m - malignant, b - benign, u - unknown) from the used datasets.

database is thus extremely fast, enabling us to compare hundreds of thousands of image feature distributions per second on an ordinary computer.

#### 4 EXPERIMENTAL SKIN DATA

We verified the proposed method on the publicly available skin image International Skin Imaging Collaboration Archive (ISIC) database (ISI ; Rotemberg et al., 2021). We used the following datasets from the

Table 1: Used datasets.

Dataset	No. of images	Image resolution [pixels]
<b>2018 JID Editorial</b>	100	various
<b>HAM10000</b>	10 015	600 × 450
<b>MSK</b>	3 918	various
<b>SONIC</b>	9 251	3024 × 2016
<b>UDA</b>	617	various
<b>total</b>	23 901	

ISIC archive (Tab. 1) to be able to compare our results with the results published in (Favole et al., 2020):

The 2018 JID Editorial dataset contains selected 100 sequentially biopsied cutaneous melanomas (37), basal cell carcinomas (40), and squamous cell carcinomas (23) with high-quality clinical images (Navarrete-Dechent et al., 2018) from the ISIC archive. All lesions originated from Caucasian patients in the southern United States. Each lesion in the skin datasets is represented by a single image in the JPEG format.

The HAM10000 dataset (Tschandl et al., 2018) is made up of 10 015 images and was collected over 20 years. Thus, the older images were digitized from slides with a Nikon Coolscan 5000 ED scanner in 300 DPI resolution. The more recent images were acquired from the digital dermatology system Mole-Max HD or the DermLiteTM Foto camera.

The MSK dataset (Codella et al., 2018) which contains five subsets made up of 3918 in total. This set with 3 to 5 images per lesion was acquired using a dermoscopic attachment to either a digital single reflex lens (SLR) camera or a smartphone (Rotemberg et al., 2021).

The SONIC dataset (Gomez and Herrera, 2017) contains 9251 images from SONIC Healthcare, Australia, which is acquired with Fujifilm FinePix S2 Pro and 176 Nikon D300 cameras.

The UDA dataset contains two subsets which are made up of 617 images. It includes melanoma and benign lesions with a histopathological diagnosis or clinically benign history containing metadata with patient age, diagnosis, gender, and anatomic location.

The ISIC data is highly unbalanced, with 80% benign, 10% malignant, and 10% unknown images, this severe underrepresentation of the most common skin lesions can lead to a much lower diagnostic accuracy when many samples are wrongly assigned to the same class. Authors (Favole et al., 2020) tried to prevent this problem by adding an altered version of malignant and unknown images. We did not use such help to see if our features can outperform the deep neural net results even with a much smaller learning data set.

Fig. 4 illustrates the malignant, benign, or unknown image examples of all five datasets. We have used the leave-one-out approach for the classification rate estimation and only on the original dataset without any data augmentation as was done in the compared deep neural net results (Favole et al., 2020). Although in (Favole et al., 2020) the ratio of training and testing sets is 80:20, due to the data augmentation, they still use the double size training set than the presented method. Thus both validation approaches (hold-out and leave-one-out) can be compared. The augmented data in their validation set ((Favole et al., 2020)) artificially increased their accuracy. Thus our improvement in Tab. 3 would be in the correct comparison even higher.

Table 2: The 2DSCAR method results for used skin datasets.

	Benign gr. truth	Malignant gr. truth	Unknow gr. truth	Precision [%]
<b>Benign int.</b>	17941	858	454	93.2
<b>Malignant int.</b>	965	1175	279	48.6
<b>Unknown int.</b>	467	252	1510	67.7
<b>Sensitivity [%]</b>	92.6	51.4	67.3	<b>Accuracy</b> 86.3

## 5 RESULTS

Our experiment compares three separate lesion classes - benign, malignant, and unknown images using a single resolution level. We have reached 86.3% accuracy on the selected dataset. The sensitivity for all classes is between 51.4 – 92.6 [%] with median value 67.3% and precision 48.6 – 93.2 [%] with median value 67.7%. More details about the results are in Tab. 2. If we added all flipped versions of malignant and unknown images as in (Favole et al., 2020), the test database would contain 52% benign, 24% malignant, and 24% unknown images, and the total number of images in the test database is 37485. The accuracy would be improved to 95% and significantly in malignant (precision 94%, sensitivity 99.7%) and unknown images (precision 88%, sensitivity 99.7%).

We compared the accuracy results of the 2DSCAR with the three CNNs evaluated in (Favole et al., 2020) and additional three recently published CNN methods. The comparative results are in Tab. 3 with each method's number of parameters. The presented method has three order fewer parameters than the alternative CNNs, which means that our method needs much less data for our reliable model learning. The training times reported in (Favole et al., 2020) are between 47 minutes for the fastest Inception-V1 until 72

Table 3: Accuracy comparison. Results denoted \* use pre-trained nets with transfer learning.

Method	No. of parameters	Accuracy	Precision	Sensitivity	Specificity
<b>2DSCAR</b>	$1.6 \cdot 10^3$	0.86	0.70	0.71	0.87
<b>2DSCAR augmented</b>	$1.6 \cdot 10^3$	0.95	0.94	0.97	0.98
AlexNet (Favole et al., 2020)	$6 \cdot 10^7$	0.74	-	-	-
Inception-V1 (Favole et al., 2020)	$5 \cdot 10^6$	0.70	-	-	-
RestNet50 (Favole et al., 2020)	$2.6 \cdot 10^7$	0.74	-	0.75	0.86
ARL-CNN50* (Zhang et al., 2019)	$2.3 \cdot 10^7$	0.86	-	0.77	0.88
R-CNN* (Khan et al., 2021)	$2.3 \cdot 10^7$	0.86	0.87	0.86	-
Hosny* (Hosny et al., 2020)	$5 \cdot 10^6$	0.98	-	0.98	0.99

minutes for RestNet50. We list the results of the last three methods (\*) for broad outline only. These methods cannot be directly compared because their authors do not specify their training data and other parameters precisely.

Although we eliminated variable illumination and the observation rotation problem using the illumination and rotationally invariant features in our study, other challenges, such as viewing angles, body position, observation distance, resolution, or variability in acquisition technologies, need to be treated for reliable and fully automatic skin diagnosis system. We plan to investigate these research problems in our future studies.

## 6 CONCLUSION

We present the method for lesion image categorization. The classifier uses rotationally invariant monospectral Markovian textural features from all three spectral classes. Our textural features are analytically derived from the underlying descriptive textural model and can be efficiently, recursively, and adaptively learned. Our 2DSCAR features are rotationally invariant, exploit information from all spectral bands, and can be easily parallelized or made fully illumination invariant if the non-illumination invariant features are excluded (the posterior probability density and the absolute error of the one-step-ahead prediction). The classifier does not need extensive learning data contrary to the convolutional neural nets and outperforms in classification accuracy three deep network nets on the same data set with more than 12%. If we add flipped versions of the malignant and unknown images for learning as was done with the alternative three deep network nets, we could outperform these methods up to 20%. Another benefit of the presented method is the order of magnitude faster learning.

## ACKNOWLEDGMENT

The Czech Science Foundation project GAČR 19-12340S supported this research.

## REFERENCES

- Isic - the international skin imaging collaboration, university dermatology center.
- Akram, T., Lodhi, H. M. J., Naqvi, S. R., Naeem, S., Alhaisoni, M., Ali, M., Haider, S. A., and Qadri, N. N. (2020). A multilevel features selection framework for skin lesion classification. *Human-centric Computing and Information Sciences*, 10(1):1–26.
- Ballerini, L., Fisher, R. B., Aldridge, B., and Rees, J. (2013). A color and texture based hierarchical k-nn approach to the classification of non-melanoma skin lesions. In *Color Medical Image Analysis*, pages 63–86. Springer.
- Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., Kallou, A., Liopyris, K., Mishra, N., Kittler, H., et al. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 168–172. IEEE.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542:115–118.
- Favole, F., Trocan, M., and Yilmaz, E. (2020). Melanoma detection using deep learning. In *International Conference on Computational Collective Intelligence*, pages 816–824. Springer.
- Gessert, N., Nielsen, M., Shaikh, M., Werner, R., and Schlaefer, A. (2020). Skin lesion classification using ensembles of multi-resolution efficientnets with meta data. *MethodsX*, 7:100864.
- Gomez, C. and Herrera, D. S. (2017). Recognition of skin melanoma through dermoscopic image analysis. In Romero, E., Lepore, N., Brieva, J., and Garca, J. D.,

- editors, *13th International Conference on Medical Information Processing and Analysis*, volume 10572, pages 326 – 336. Int. Society for Optics and Photonics, SPIE.
- Haindl, M. (2012). Visual data recognition and modeling based on local markovian models. In Florack, L., Duits, R., Jongbloed, G., Lieshout, M.-C., and Davies, L., editors, *Mathematical Methods for Signal and Image Analysis and Representation*, volume 41 of *Computational Imaging and Vision*, chapter 14, pages 241–259. Springer London. 10.1007/978-1-4471-2353-8.14.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Hosny, K. M., Kassem, M. A., and Foad, M. M. (2020). Skin melanoma classification using roi and data augmentation with deep convolutional neural networks. *Multimedia Tools and Applications*, 79(33):24029–24055.
- Kawahara, J., BenTaieb, A., and Hamarneh, G. (2016). Deep features to classify skin lesions. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1397–1400. IEEE.
- Kawahara, J. and Hamarneh, G. (2019). Visual diagnosis of dermatological disorders: Human and machine performance.
- Khan, M. A., Zhang, Y.-D., Sharif, M., and Akram, T. (2021). Pixels to classes: intelligent learning framework for multiclass skin lesion localization and classification. *Computers & Electrical Engineering*, 90:106956.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, volume 25, pages 1097–1105.
- Mahbod, A., Schaefer, G., Ellinger, I., Ecker, R., Pitiot, A., and Wang, C. (2019). Fusing fine-tuned deep features for skin lesion classification. *Computerized Medical Imaging and Graphics*, 71:19–29.
- Navarrete-Dechent, C., Dusza, S. W., Liopyris, K., Marghoob, A. A., Halpern, A. C., and Marchetti, M. A. (2018). Automated dermatological diagnosis: hype or reality? *The Journal of investigative dermatology*, 138(10):2277.
- Nida, N., Irtaza, A., Javed, A., Yousaf, M. H., and Mahmood, M. T. (2019). Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering. *International journal of medical informatics*, 124:37–48.
- Remeš, V. and Haindl, M. (2018). Rotationally invariant bark recognition. In X., B., E., H., T., H., R., W., B., B., and A., R.-K., editors, *IAPR Joint International Workshop on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition (S+SSPR 2018)*, volume 11004 of *Lecture Notes in Computer Science*, pages 22 – 31. Springer Nature Switzerland AG.
- Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., et al. (2021). A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific data*, 8(1):1–8.
- Shakourian Ghalejoogh, G., Montazery Kordy, H., and Ebrahimi, F. (2019). A hierarchical structure based on stacking approach for skin lesion classification. *Expert Systems with Applications*, 145:113127.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceeding of the IEEE conference on computer vision and pattern recognition*, pages 1–9.
- Tschandl, P., Rosendahl, C., and Kittler, H. (2018). The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9.
- Tschandl, P., Sinz, C., and Kittler, H. (2019). Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. *Computers in biology and medicine*, 104:111–116.
- Zhang, J., Xie, Y., Xia, Y., and Shen, C. (2019). Attention residual learning for skin lesion classification. *IEEE transactions on medical imaging*, 38(9):2092–2103.
- Zhang, N., Cai, Y.-X., Wang, Y.-Y., Tian, Y.-T., Wang, X.-L., and Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artificial intelligence in medicine*, 102:101756.