



A Convolutional Neural Network for Lentigo Diagnosis

Sana Zorgui^{1(✉)}, Siwar Chaabene^{1(✉)}, Bassem Bouaziz^{1(✉)}, Hadj Batatia^{2(✉)},
and Lotfi Chaari^{2(✉)}

¹ MIRACL and CRNS, University of Sfax, Sfax, Tunisia
sanazorgui@gmail.com, siwarchaabene@gmail.com,
bassem.bouaziz@isims.usf.tn

² University of Toulouse, IRIT - INP-ENSEEIH, Toulouse, France
{hadj.batatia,lotfi.chaari}@toulouse-inp.fr

Abstract. Using Reflectance Confocal Microscopy (RCM) for lentigo diagnosis is today considered essential. Indeed, RCM allows fast data acquisition with a high spatial resolution of the skin. In this paper, we use a deep convolutional neural network (CNN) to perform RCM image classification in order to detect lentigo. The proposed method relies on an InceptionV3 architecture combined with data augmentation and transfer learning. The method is validated on RCM data and shows very efficient detection performance with more than 98% of accuracy.

Keywords: Reflectance Confocal Microscopy · Lentigo · CNN classification · InceptionV3

1 Introduction

Reflectance Confocal Microscopy (RCM) [1] is a modality increasingly used in medical imaging like MRI (Magnetic Resonance Imaging) [2–4] or X-ray imaging [5]. In vivo RCM technique is easy to use during the patient examination and acquires high resolution skin images in a short time. This modality can be used to help dermatologists diagnose different skin diseases. However, it takes a long time for dermatologists to make full use of the possibilities of this technique for diagnostic purposes. Our work aims to develop a new tool to automate certain diagnostic steps required using deep learning [6]. On the other side, the lentigos are age spots that mainly appear on the hand or on the areas most frequently exposed to the sunlight. On the surface, they appear as a darker spot. Inside the skin layers, it is mainly at the level of the dermis-epidermis junction that the differences can be visible [7]. Therefore, the distinction of lentigos can be made using the RCM images. Several deep learning architectures, especially convolutional neural network (CNN) [5, 8] show great potential in medical imaging classification. In this paper, we propose a new 3D RCM image (2D + depth) classification method for lentigo detection. The method is based on a CNN on InceptionV3 architecture [9].

Until now, little works have been proposed for lentigo/healthy classification of RCM images. In [10], the authors perform a two-dimensional wavelet decomposition. Then a generalized Gaussian distribution was applied to the wavelet coefficients in order to perform a quantitative analysis assisted by a support vector machine (SVM) to classify RCM images obtaining an accuracy of 84.4%. Another approach in [11] explores a new unsupervised Bayesian algorithm for the joint reconstruction and classification of RCM images. The resulting algorithm for healthy and lentigo classification reached an accuracy percentage of 97%. Beside, the paper [12] automatically diagnosed lentigo by using three separate feature extraction methods like Wavelets, Haralick and CNN by Transfer Learning. The healthy/lentigo classification results reached an accuracy of 76%.

The present paper is organized as follows. Section 2 presents the problem formulation of lentigo diagnosis. Section 3 detailed the proposed lentigo detection method. Section 4 presents the experiment validation of our method. Finally, conclusion and some perspectives are drawn in Sect. 5.

2 Related Work

2.1 Lentigo Detection

Lentigo is a lesion that occurs in the dermal epidermal junction between the dermis and the epidermis involving a high concentration of melanocytes in the dermal papillae walls. Most forms of lentigo are benign [13] like lentigo simplex as Fig. 1(a) and solar lentigo as Fig. 1(b). They are usually removed for cosmetic purposes. However, certain types such as lentigo maligna [14] as Fig. 1(c) may be harmful and must be removed.



Fig. 1. Lentigo simplex (a), Solar lentigo (b) and Lentigo maligna (c).

Usually, lentigo is diagnosed using dermatoscopy [15]. Sadly, non-pigmented melanocytes with this modality can go completely unnoticed leading to complications in identifying the lesion contours with precision. Histopathology [16] is also used to confirm the diagnosis, but it can be inconvenient due to the fact that it is an in vitro technique involving performing a biopsy from the pigmented areas. For these reasons, the RCM modality emerged to solve the problems encountered before. Therefore, this modality allows the expert to carry out a real-time 3D

data acquisition and to facilitate the full observation of the biological structures in deformation over time. Due to all of these reasons, our approach is based on images acquired thanks to this modality. In [10, 11], the authors propose two RCM lentigo detection methods based on the statistical and Bayesian models [17] respectively. The methods have proved complicated and hard to implement. They require manual procedures like feature selection and data preparation. To this regard, we propose here a method for RCM image classification using a CNN architecture. Indeed, CNNs have proven their capacity to efficiently solve several complex problems in medical imaging.

2.2 Convolutional Neural Networks

The CNN [8] is a deep learning architecture that is primarily used for image classification and object detection. Figure 2 displays a general CNN architecture, where one can easily identify the following layers:

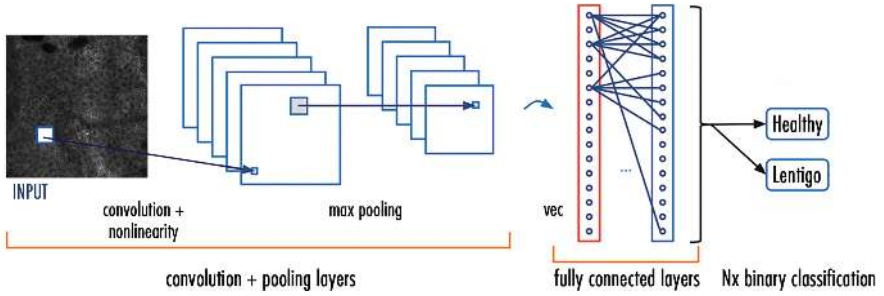


Fig. 2. The CNN architecture model.

- The convolutional layers: a key component of a CNN architecture, used for automatic feature extraction.
- The rectified linear units (ReLU): used after each convolutional layer. Each layer combines nonlinear layers and rectification layers to add nonlinearity to the system.
- The pooling layers: used for feature selection by maximum or/and average pooling.
- The fully connected layers: also known as dense layers receiving the flattened (1D) feature map. Usually, the final fully connected layer has the same number of output nodes as the class numbers.
- The Softmax function: calculates the probabilities of each target class over all possible classes. This function helps determine the target class for the given inputs.

3 Proposed Method for Lentigo Detection

The proposed method consists of classifying RCM images into healthy/lentigo classes using an InceptionV3 architecture. Our lentigo detection method combines the InceptionV3 model with other known deep learning techniques like transfer learning [18] and data augmentation [19]. Figure 3 presents the different steps used in the proposed lentigo detection method. The following subsections give detailed descriptions of each step.

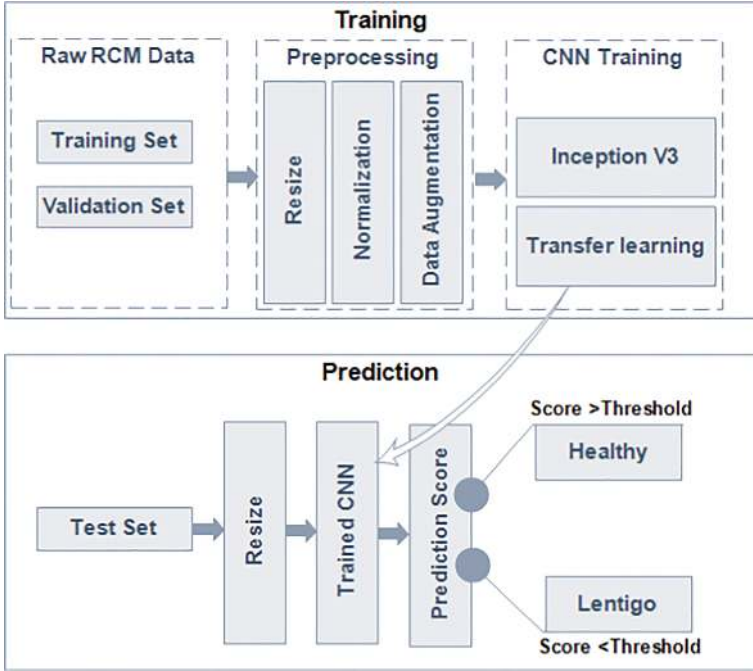


Fig. 3. Pipeline of the proposed method.

3.1 Data Preparation

The input RCM images for the training procedure combines two sets such as a 73% training set and a 14% validation set. The remaining 13% is dedicated to the prediction phase. In order to avoid overfitting, a validation set is added to our training phase because the non linear InceptionV3 model will possibly achieve 100% training accuracy and overfit.

3.2 Data Preprocessing

In the first step of the preprocessing procedure, the RCM images of the training set are resized to fit in the InceptionV3 network. A normalization step is added to help the CNN better process the input images, in order that all feature values have the same range and the system needs only one global learning rate multiplier. Afterwards, the data augmentation step is proposed to improve our classification results. This step prevents accuracy decay and overfitting. In [20] the authors demonstrate the importance of data augmentation as a regularizer in the CNN classification model.

3.3 InceptionV3 Model

The InceptionV3 model is a complex heavily engineered network that considered a major breakthrough in CNN’s [9]. Before the current model, many common CNN’s claimed that stacking layers after layers is the only way to increase accuracy. However, this network suggested some solutions to improve accuracy and speed without piling many layers. As shown in Fig. 4, the InceptionV3 model consists of a combination of three main modules.

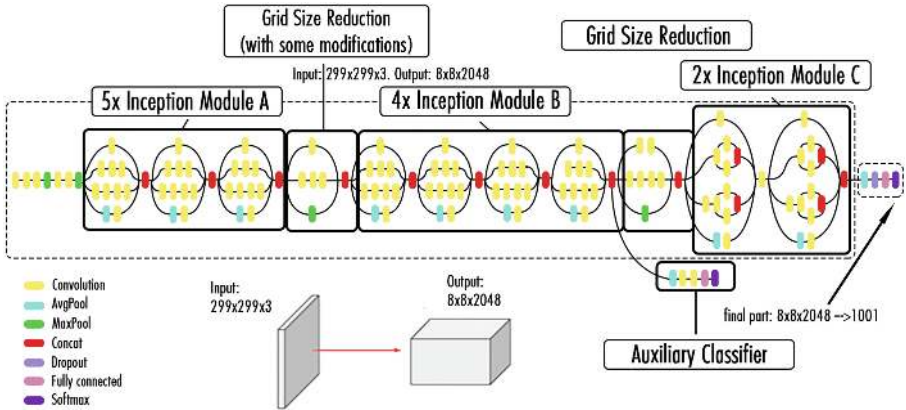


Fig. 4. Architecture of the InceptionV3 model.

The first one (Module A) uses two smaller convolution layers (3×3) to decrease the computational cost by reducing the number of parameters to improve performance. Module B divides each convolution layer of $n \times n$ size to two layers of $1 \times n$ and $n \times 1$ dimensions to have a less complex network. Finally, Module C reduces the representational bottleneck by expanding the filters in order to evade information loss. More upgrades are also proposed by the InceptionV3 network other than the smart factorization methods such as:

- RMSProp optimizer allows a faster convergence of the model thus allowing a higher learning rate.

- BatchNorm reduces the covariance shift and allows each network layer to learn a little independently of the others.
- Label Smoothing is a regularizing component applied to the loss formula to prevent overfitting.

The InceptionV3 network is 42 layers deep. Therefore, the computational cost is just around 2.5 higher than GoogLeNet’s [21]. In addition, the inception modules are a novel and popular concept due to their smaller convolutions, which explains the reduction in the number of parameters. The InceptionV3 model gathers more information without impacting the computational speed thanks to its depth and the various kernel sizes used in the convolution operations.

3.4 Transfer Learning

As shown in Fig. 3, transfer learning [22] is proposed in order to ensure better performance of the model. The model needs lots of labeled images to be capable of solving complex problems. This has proved to be challenging especially when the available dataset is small. Transfer learning is a deep learning method, in which a model developed for a task is reused for a second task. This technique uses pre-trained models as a starting point for other medical imaging tasks given the vast computational and time resources required to develop CNN models on these problems.

3.5 Prediction Model

In the prediction phase, the RCM images test set are resized and provided to the trained CNN. Our system calculates a prediction score for each test image after resizing it and compares it with the threshold T equal to 0.5. The threshold value is chosen that way due to the fact that we are performing a binary classification. The classification condition is as follows: if the predicted score (PS) value of the image test is lower than T then this RCM image is classified as lentigo and conversely.

4 Experimental Validation

This section evaluates the validation of the proposed lentigo detection method on real RCM data. In our work, the dataset is provided from Lab. Pierre Fabre. In this experiment, the data include 428 RCM images which high spatial resolutions and annotation on each image into two healthy and lentigo classes. The images were acquired with a Vivascope 1500 apparatus. Each RCM image shows a field of view of $500 \times 500 \mu\text{m}$ with 1000×1000 pixels. A selection of 45 women aged 60 years were recruited. All participants have offered their informed consent to the RCM skin test. We split these data into three main sets:

- A 314 images training set divided into two classes of 160 healthy images and 154 lentigo images.

- A validation set of 60 images, divided equally between two classes of lentigo and healthy. The validation set has been added to evaluate our training procedure. The main objective is to prevent over-fitting.
- A 54 RCM images testing set divided equally for healthy and lentigo classes.

Our classification method based on the InceptionV3 network is build using the Keras library. The InceptionV3 model is configured to accept the greyscale RCM images. As initialization, all RCM images were resized into new dimensions of 299×299 pixels and rescaled to help CNN processing. The parameter values of data augmentation step are presented in Table 1. The shear, zoom and translation ranges vary from 0 to 1. We choose the value of 0.2 for each to enrich the dataset without altering the image main features and confusing the system. The rotation range varies to 0° from 180° and a small rotation angle was proposed for the same reasons.

Table 1. Data augmentation parameters.

| Parameter | value |
|------------------------|------------|
| Shear | 0.2 |
| Zoom | 0.2 |
| Rotation degree | 20° |
| Horizontal translation | 0.2 |
| Vertical translation | 0.2 |

Figure 5 displays the accuracy curves of the training and validation sets, as well as the training loss. The accuracy curves suggest that our system converged after 40 epochs. The system reached an accuracy value of 94% for training and 69% for validation. Hence, the reported values indicate that our system learns well without over- or under-fitting.

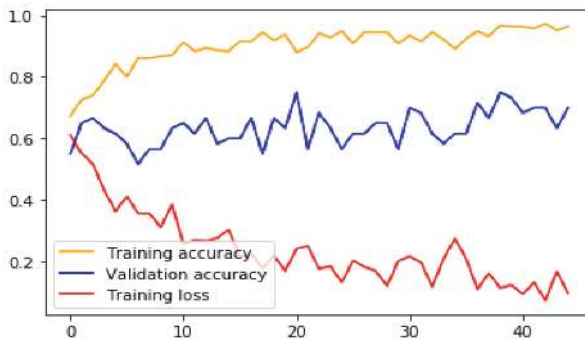


Fig. 5. Proposed method accuracy graph and loss graph for training and validation sets after each epoch.

The performance of the proposed method is indicated by the test set according to the ability to correctly diagnose the provided skin tissues. The reported values in Table 2 indicate the performance of our classification method. Therefore, 53 out of 54 images test set were correctly classified with an accuracy of **98,14%**.

Table 2. Confusion matrix.

| | Lentigo | Sane |
|---------|-------------------|--------------------|
| Lentigo | 27/27 = 100% (TP) | 1/27 = 3,7% (FN) |
| Sane | 0/27 = 0% (FP) | 26/27 = 96,3% (TN) |

In Table 2, TP, TN, FP and FN represent respectively true positives, true negatives, false positives and false negatives. Based on the confusion matrix, Accuracy, Precision, Specificity, Recall and F-score values are reported in Table 3. All the mentioned measures indicate a good performance of the proposed method with values equal or very close to one.

Table 3. Quantitative evaluation of the proposed method performance.

| | | |
|----------------------|---------------------------------|------|
| Accuracy | $(TP + TN)/(TP + TN + FP + FN)$ | 0.98 |
| Precision (P) | $TP/(TP + FP)$ | 1 |
| Specificity | $TN/(FP + TN)$ | 1 |
| Recall (R) | $TP/(TP + FN)$ | 0.96 |
| F-score | $(2 \times P \times R)/(P + R)$ | 0.97 |

Figure 6 presents four correct classification examples of RCM images from the test set. The reported values shown with each test image indicate the prediction score (PS). The displayed images correspond to different PS ranges. We can notice that the model performs well both for images with PS close to 0 or 1, but also for images with PS close to 0.5 (images (b) and (d)).

Figure 7 shows the only image wrongly classified using our proposed method. This image shows some type of skin deformation similar to the changes the skin undergoes due to lentigo. Hence, the network interpreted it as a lentigo lesion. For the sake of further evaluation, we compare the accuracy of the test with related works that used the same dataset. The reported values in Table 4 show that our model outperforms in comparison with the other methods. Specifically, we compare our results with those reported in [10] where the authors used a Statistical model combined with an SVM classifier and [11] where the authors use an unsupervised Bayesian approach.

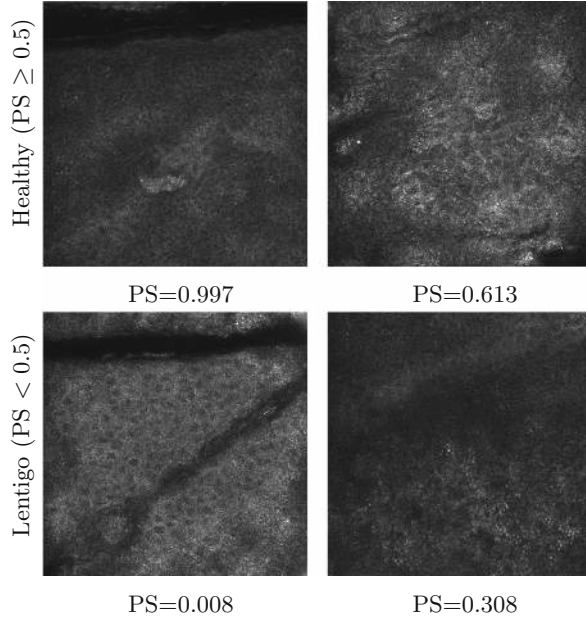


Fig. 6. Correct classification examples of RCM images for Healthy and Lentigo patients classified by the proposed method.

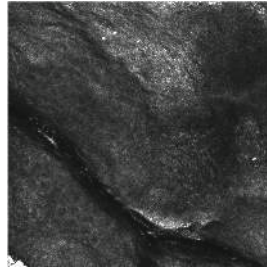


Fig. 7. The only false classification ($PS = 0.0005$).

Table 4. Comparison performance with state of the art methods.

| Lentigo detection method | Accuracy |
|--------------------------|---------------|
| Halimi et al. 2017 [10] | 84.4% |
| Halimi et al. 2017 [11] | 97.7% |
| Proposed method | 98,14% |

5 Conclusion

In this paper, we proposed a new method to classify RCM images into healthy and lentigo skins. This method is based on the InceptionV3 CNN architecture. The network was trained with a dataset of 374 images and tested on 54 images of different stacks and depths. The suggested CNN method shows huge potential and very promising results. In future work, we will focus on applying the proposed approach to larger datasets and comparisons to other deep architectures.

Acknowledgements. The authors would like to thank Gwendal JOSSE et Jimmy Le Digabel from Lab. Pierre Fabre for providing data.

References

1. Rajadhyaksha, M., Grossman, M., Esterowitz, D., Webb, R.H., Anderson, R.R.: In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. *J. Invest. Dermatol.* **104**, 946–952 (1995)
2. Laruelo, A., et al.: Hybrid sparse regularization for magnetic resonance spectroscopy. In: IEEE International Conference of Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013, pp. 6768–6771 (2013)
3. Albughdadi, M., Chaari, L., Tourneret, J.Y., Forbes, F., Ciuciu, P.: A Bayesian non-parametric hidden Markov random model for hemodynamic brain parcellation. *Sig. Process.* **135**(10223), 132–146 (2017)
4. Chaabene, S., Chaari, L., Kallel, A.: Bayesian sparse regularization for parallel MRI reconstruction using complex Bernoulli–Laplace mixture priors. *SIViP* **14**(3), 445–453 (2019). <https://doi.org/10.1007/s11760-019-01567-5>
5. Fakhfakh, M., Bouaziz, B., Gargouri, F., Chaari, L.: 1ProgNet: Covid-19 prognosis using recurrent and convolutional neural networks. *IEEE Trans. Artif. Intell.* (2020, submitted)
6. Geert, L., et al.: A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017)
7. Calzavara-Pinton, P., Longo, C., Venturini, M., Sala, R., Pellacani, G.: Reflectance confocal microscopy for in vivo skin imaging. *Photochem. Photobiol.* **84**, 1421–1430 (2008)
8. Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K.: Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**(4), 611–629 (2018). <https://doi.org/10.1007/s13244-018-0639-9>
9. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. *Computing Research Repository (CoRR)* (2015)
10. Halimi, A., Batatia, H., Digabel, J., Josse, G., Tourneret, J.Y.: Statistical modeling and classification of reflectance confocal microscopy images. In: *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 1–5, December 2017
11. Halimi, A., Batatia, H., Digabel, J., Josse, G., Tourneret, J.Y.: An unsupervised Bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images. In: *European Signal Processing Conference EUSIPCO*, pp. 241–245, August 2017

12. Cendre, R., Mansouri, A., Benezeth, Y., Marzani, F., Jean, P., Cinotti, E.: Two schemes for automated diagnosis of lentigo on confocal microscopy images. In: International Conference on Signal and Image Processing (ICSIP), pp. 143–147, July 2019
13. Ève, O.: Les produits dépigmentants: le point en 2011, p. 78, September 2011
14. Cohen, L.M.: Lentigo maligna and lentigo maligna melanoma. *J. Am. Acad. Dermatol.* **33**(6), 923–936 (1995)
15. Bollea-Garlatti, L.A., Galimberti, G.N., Galimberti, R.L.: Lentigo maligna: keys to dermoscopic diagnosis. *Actas Dermo-Sifiliográficas (English Edition)* **107**(6), 489–497 (2016)
16. Andersen, W.K., Labadie, R.R., Bhawan, J.: Histopathology of solar lentiginos of the face: a quantitative study. *J. Am. Acad. Dermatol.* **36**(3), 444–447 (1997)
17. Chaari, L.: A Bayesian grouplet transform. *SIViP* **13**(5), 871–878 (2019). <https://doi.org/10.1007/s11760-019-01423-6>
18. Ribani, R., Marengoni, M.: A survey of transfer learning for convolutional neural networks. In: Conference on Graphics, Patterns and Images Tutorials, pp. 47–57 (2019)
19. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**(1), 1–48 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
20. García, A.H., König, P.: Further advantages of data augmentation on convolutional neural networks. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds.) ICANN 2018. LNCS, vol. 11139, pp. 95–103. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01418-6_10
21. Szegedy, C., et al.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, June 2015
22. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? Computing Research Repository (CoRR), pp. 3320–3328 (2014)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

