DESDE 2013
https://repository.uaeh.edu.mx/revistas/index.php/icbi/issue/archive
Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI
ISSN: 2007-6363

Publicación Semestral Pädi Vol. 10 No. Especial 3 (2022) 1–4

Rotation-invariant image classification using a novel 1D CNN and Multichannel Accurate Bessel-Fourier moments
Clasificación de imágenes invariantes a la rotación utilizando una novedosa CNN en 1D y Momentos exactos de Bessel-Fourier

C. J. Camacho-Bello [a,*], L. Gutiérrez-Lazcano [a], R. M. Ortega-Mendoza [a]

[a] *Dirección de Investigación y Posgrado, Universidad Politécnica de Tulancingo, 43629, Tulancingo, Hidalgo, México.*

Resumen

Este trabajo presenta una propuesta para utilizar momentos de Bessel-Fourier como entradas a redes neuronales convolucionales 1D de tal manera que aprovechen las características inherentes de los descriptores de tipo de momento, como la invariancia rotacional y la mínima redundancia de información. Los resultados presentados muestran que la propuesta tiene un mejor desempeño que la red neuronal profunda con invariancia de rotación.

Palabras Clave: Momentos de Bessel-Fourier, características de invariante rotación , aprendizaje profundo, morfologías de galaxias.

Abstract

This work presents a proposal to use Bessel-Fourier moments as inputs to 1D convolutional neural networks in such a way that they take advantage of the inherent characteristics of moment type descriptors such as rotational invariance and minimal information redundancy. The results presented show that the proposal has a better performance than the deep neural network with rotation invariance.

Keywords: Bessel-Fourier moments, Rotation-Invariant features, Deep learning, Galaxy morphologies.

1. Introduction

Orthogonal moments are rotational invariant image descriptors with minimal information redundancy (Teague, 1980). Descriptor-based classifiers differ primarily in the characteristics that local or global. Generally, the most used methods are distance and predefined similarity measures. However, only a few studies have considered 1D convolutional neural networks to accomplish this task; this is surprising since deep learning has revolutionized different machine learning fields by developing novel neural networks.

In general, the problem related to classifying images affected by random orientations can be treated differently. One way to tackle this problem is to integrate rotated samples into the training dataset (Cheng et al., 2016). However, this does not help to demonstrate that features are rotation-invariant because the classifier does not consider each object's shape features. Another approach is based on extracting rotational invariant descriptors such as an object's descriptors will be identical, regardless of the object's orientation.

On the other hand, several advances in endowing a deep neural network with rotation invariance by modifying the network's internal filters to obtain characteristics invariant to rotation; the approach includes Oriented Response Networks (ORN) (Zhou et al., 2017), Covariant CNN (Rodriguez et al., 2019), Rotation Invariant Convolutional (RIC) (Follmann and Bottger, 2018), and Rotation Invariant Networks (RIN) (Rodriguez Salas et al., 2021). Also, few studies train their algorithms with upright images database and validate rotated samples (Zhou et al. (2017); Follmann and Bottger (2018); Rodriguez et al. (2019); Rodriguez Salas et al. (2021)). In state-of-the-art, invariant descriptors' performance is validated using the MNIST database of handwritten digit images to train and the MNIST-R database to validate the results.

For this reason, we propose to replace the modified filters with moment type descriptors such as Bessel-Fourier moments.

*Autor para correspondencia: cesar.camacho@upt.edu.mx

Correo electrónico: cesar.camacho@upt.edu.mx (César Joel Camacho-Bello), lucia.gutierrez@upt.edu.mx (Lucia Gutiérrez-Lazcano), rosa.ortega@upt.edu.mx (Rosa María Ortega-Mendoza).

Historial del manuscrito: recibido el 28/03/2022, última versión-revisada recibida el 08/07/2022, aceptado el 02/08/2022, publicado el 31/08/2022. **DOI:** https://doi.org/10.29057/icbi.v10iEspecial3.8874

For classification, we suggest a variation of the fully convolutional neural network presented by Wang et al. (2017), which performs well for time series classification.

In addition, we compare different deep neural network models in a galaxy morphology classification. The proposal's performance is also explored since the galaxies' rotation should not affect their type taken as a reference the work of Martinazzo et al. (2020), where they use the EFIGI database (De Lapparent et al., 2011) and perform the classification using the DenseNet121 (Huang et al., 2017), ResNext50 (Xie et al., 2017), VGG16 (Simonyan and Zisserman, 2014), and InceptionV3 (Szegedy et al., 2016) with learning transfer and parameter tuning. Moreover, a brief description of the Bessel-Fourier moments is presented; later, the proposed model is compared with the state-of-the-art dataset to test the rotation invariance and compare with different network models to classify galaxies. Finally, the conclusions are presented.

2. Bessel-Fourier moments

Bin Xiao et al. (2010) proposes the Bessel-Fourier moments, which use the first kind Bessel function as kernel, they have the ability to characterize, evaluate, and manipulate visual information with minimum redundant information. The Bessel-Fourier moments of order n with repetition m for an image in polar coordinates $f(r,\theta)$ are defined as,

$$B_{n,m} = \frac{1}{2\pi a_n} \int_0^{2\pi} \int_0^1 f(r,\theta) J_1(\lambda_n r) \exp(-jm\theta) r dr d\theta \quad (1)$$

where $a_n = [J_2(\lambda_n)]^2/2$ is the normalization constant, $J_1(\lambda_n r)$ is the Bessel function, and λ_n is the nth zero of the Bessel function. On the other hand, to represent a color image, Singh and Singh (2018) defined multi-channel moments as a set of moments for each channel of a color image $f_p(r,\theta)$ with three primary channels $f_R(r,\theta)$, $f_G(r,\theta)$, and $f_B(r,\theta)$. The multi-channel Bessel-Fourier moments can be represented as,

$$B^p_{n,m} = \frac{1}{2\pi a_n} \int_0^{2\pi} \int_0^1 f_p(r,\theta) J_1(\lambda_n r) \exp(-jm\theta) r dr d\theta, \quad (2)$$

where $p = \{R, G, B\}$ and represent the channels of the RGB color space.

On the other hand, image reconstruction is a tool to identify the number of descriptors or moments needed to describe an image. The reconstructed discrete distribution of the image is given by

$$\widetilde{f}(i,j) = \sum_{n=0}^{L} \sum_{m=-n}^{n} B^p_{n,m} J_1(\lambda_n r) \exp(-jm\theta) \quad (3)$$

where $\widetilde{f}(i,j)$ is the reconstructed version of $f(i,j)$, and L is the maximum order. Fig. 1 shows the reconstruction of an image from the MNIST database.

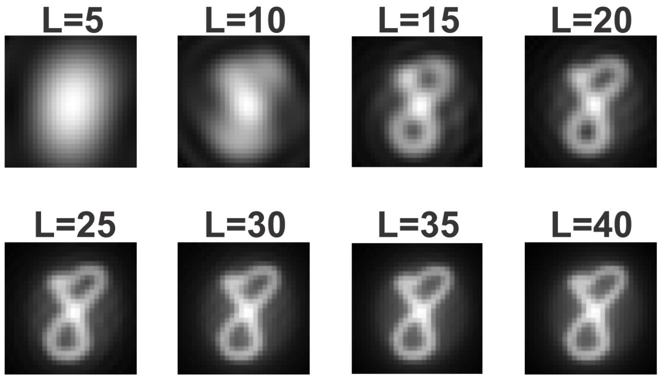


Figura 1: Image Reconstruction: from the order $L > 15$ the image can be recovered.

One of the most important properties of the orthogonal moments defined in a unit disk is its invariant representation of the rotation. If we consider an image $f(r,\theta-\gamma)$ is rotated γ degrees, then $B^{(\gamma)}_{n,m}$ are related with the Bessel-Fourier moments of the original image by

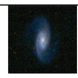
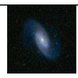
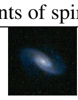
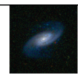
$$B^{(\theta-\gamma)}_{n,m} = B^p_{n,m} \exp(-im\gamma). \quad (4)$$

Therefore, when calculating the modulus of Bessel-Fourier moments,

$$\left| B^{(\gamma)}_{n,m} \right| = \left| B^p_{n,m} \right|, \quad (5)$$

it verifies that the circular moments are invariant to the rotation. Table 1 shows an example of the rotation invariance that helps in the galaxies classification.

Tabla 1: Bessel-Fourier moments of spiral galaxy with different rotation

					
B^R_{00}	5.65	5.57	5.57	5.65	5.57
B^R_{01}	0.11	0.12	0.12	0.11	0.12
B^G_{00}	7.83	7.78	7.78	7.83	7.78
B^G_{01}	0.15	0.15	0.15	0.15	0.15
B^B_{00}	8.61	8.55	8.55	8.61	8.55
B^B_{01}	0.21	0.21	0.21	0.21	0.21

3. Rotation-invariant Image classification

Bessel-Fourier moments have relevant features to represent images with a minimum amount of information; this quality is beneficial when data augmentation or pre-processing is required to solve the problem of changes in scale and rotation. On the other hand, combining moment-type descriptors and deep neural networks has been little explored because there are different elements to consider that make their design difficult.
They are currently using convolutional neural networks (CNN) and genetic algorithms (GA) to help CNN learn better and more efficiently. Some required values cannot be learned, such as the number of layers, nodes, size filters, and activation functions. In this work, the optimization of the neural network design using GA seeks to use radial moments as inputs to 1D convolutional neural networks to improve the recognition performance of

rotation-invariance classification. We are taking as a base model the one proposed by Zhiguang Wang et al. (2017) to generate the initial population. The GA helped determine the number of convolution layers, the size of the convolution kernel, the activation function, the number of dense layer nodes, and the pooling and dropout operators.

The table 2 details the proposal of the specific parameters of each layer. Also, Fig. 2 shows a scheme to represent the process of the proposed architecture.

Tabla 2: Details of each layer's parameters of the proposed 1D CNN model.

Layer	Type	Output	Kernel	Activation
1	BFM	64	-	-
2	Convolution 1D	64×25	40	relu
3	BatchNormalization	64×25	-	-
4	Convolution 1D	64×25	40	relu
5	BatchNormalization	64×25	-	-
6	Convolution 1D	64×25	40	relu
7	BatchNormalization	64×25	-	-
8	Convolution 1D	64×25	40	relu
9	BatchNormalization	64×25	-	-
10	Convolution 1D	64×180	25	relu
11	BatchNormalization	64×180	-	-
12	Max-pooling 1D	35×180	2	-
13	Flatten	17500	-	relu
14	Dense	500	-	-
15	Dense	10	-	softmax

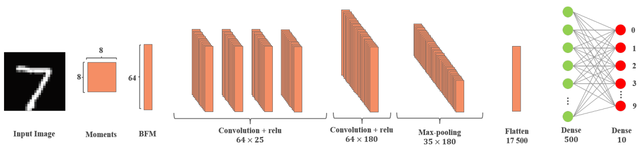


Figura 2: Scheme of the proposed architecture

To demonstrate the rotation invariant of the proposed method, we use the existing implementations in state-of-the-art results on rotation-invariance through the training MNIST database (LeCun et al., 2010) (60,000 images) and tested with MNIST-rot (Larochelle et al., 2007) (50,000 images). Figure 1 shows examples of datasets with up-right and randomly rotated images.

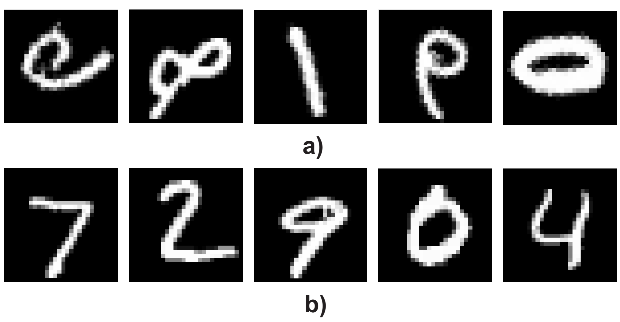


Figura 3: Examples of databases: MNIST-Rotation a) and MNIST b)

On the other hand, given the rotation invariants' characteristics, the Bessel-Fourier moments can help to determine galaxies' morphology. The EFIGI catalog (De Lapparent et al., 2011) contains 4458 galaxies, which can demonstrate the efficiency of the proposed method with different CNN. Figure 2 shows samples from the different types of galaxy images used in this study. In Martinazzo et al. (2020), they divide the morphology of galaxies into easy (EF-2), medium (EF-4), and challenging (EF-15). Each dataset is detailed:

- **EF-2:** 3604 images divided into two classes: Elliptical (289) and Spiral (3315) galaxies
- **EF-4:** 4389 images divided into four classes: Elliptical (289), Spiral (3315), Lenticular (537) and Irregular (248) galaxies
- **EF-15**: 4327 images divided into fifteen classes: Elliptical:-5 (227), Spiral:0 (196), Spiral:1 (257), Spiral:2 (219), Spiral: 3 (517), Spiral:4 (472), Spiral:5 (303), Spiral:6 (448), Spiral:7 (285), Spiral:8 (355), Spiral:9 (263), Lenticular:-3 (189), Lenticular:-2 (196), Lenticular:-1 (152), and Irregular:10 (248) galaxies.

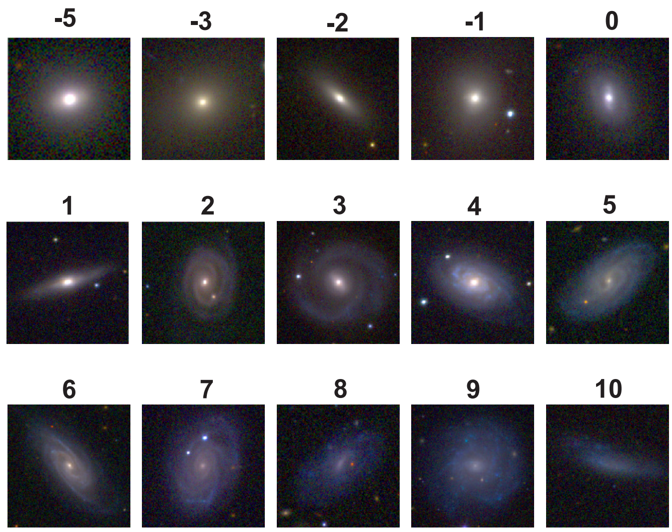


Figura 4: Example of galaxy images from the EFIGI dataset with morphological types.

The EFIGI dataset is split randomly into two: training set (90 %) and testing set (10 %). The training set is used in the learning phase, while the test set is used to evaluate the proposed model in terms of accuracy. The results of the proposed approach to classify the two datasets are given in Tables 3 and 4 where the obtained results ensure the superiority of the proposed method over the existing state-of-the-art. All experiments were executed on a computer with an Intel Core i7-9700 running at 3.0 GHz, 32 GB of RAM, and an NVidia GeForce GTX 1660 SUPER GPU with 6 GB of Video RAM. The proposal is implemented in Python with the TensorFlow libraries. Training can be GPU or non-GPU because it only has 160,000 trainable variables, unlike other architectures with millions.

Tabla 3: Error rate with training MNIST and validated with MNIST-R.

Method	Error rate
ORN-8(ORPooling) Zhou et al. (2017)	16.67 %
ORN-8(ORAlign) Zhou et al. (2017)	16.21 %
RIC(RP RF 1) Follmann and Bottger (2018)	19.85 %
RIC(RP RF 1 32) Follmann and Bottger (2018)	12.20 %
Covariant CNN Rodriguez et al. (2019)	17.21 %
RIN(G+C) Rodriguez Salas et al. (2021)	1.7 %
BFM-CNN (Proposal)	1.4 %

Tabla 4: Results of the models implemented by Martinazzo et. al Martinazzo et al. (2020) and the proposal.

	DenseNet	Inception	ResNext	VGG16	Proposal
EF-15	40.8 %	30.3 %	33.3 %	45.9 %	42.2 %
EF-4	88.1 %	83.0 %	84.1 %	88.3 %	90.4 %
EF-2	99.4 %	99.2 %	98.6 %	99.4 %	99.7 %

4. Concluding remarks

This work proposes a novel but straightforward 1D convolutional neural network to image classification degraded by random rotation with the Bessel-Fourier moments, surpassing other deep neural networks to image classification exposed to random rotations. On the other hand, databases that take advantage of the Bessel-Fourier moments' invariant characteristics were selected to demonstrate their potential as image descriptors. Undoubtedly, recent advances in neural networks significantly outperform moment-type descriptors. However, the proposed descriptors can be considered since they do not require data augmenting to extract characteristics invariant to rotation and facilitate their training. The proposed model has a similar performance to the deep learning models, such as those shown in Tables 3 and 4. To make it, the dataset to be classified must meet the following characteristics: (1) well-defined shapes, (2) centered, and (3) without background.

The proposal is the first approach to include moment-type descriptors with deep neural networks. However, there is still a long way to improve the results; it is necessary to explore the logic of other classification architectures such as ResNet, DenseNet, or Transformers. In addition, to explore different search algorithms for neural network design, such as the artificial bee colony (ABC) or the firefly algorithm (FA), which could have better results than GA.

On the other hand, the significant contribution to achieving a better classification with 1D convolutional neural networks is that they use large filters due to the minimal redundancy of the proposed neural network's input values. Furthermore, an advantage is that it can be adapted to other families of circular orthogonal moments. The proposal is easy for real-world application development and presents a starting point for future research.

Referencias

Cheng, G., Zhou, P., and Han, J. (2016). Learning rotation-invariant convolutional neural networks for object detection in vhr optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12):7405–7415.

De Lapparent, V., Baillard, A., and Bertin, E. (2011). The efigi catalogue of 4458 nearby galaxies with morphology-ii. statistical properties along the hubble sequence. *Astronomy & Astrophysics*, 532:A75.

Follmann, P. and Bottger, T. (2018). A rotationally-invariant convolution module by feature map back-rotation. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 784–792.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

Larochelle, H., Erhan, D., Courville, A., Bergstra, J., and Bengio, Y. (2007). An empirical evaluation of deep architectures on problems with many factors of variation. In *Proceedings of the 24th International Conference on Machine Learning*, ICML ?07, page 473?480, New York, NY, USA. Association for Computing Machinery.

LeCun, Y., Cortes, C., and Burges, C. J. (2010). Mnist handwritten digit database. 2010.

Martinazzo, A., Espadoto, M., and Hirata, N. S. (2020). Deep learning for astronomical object classification: A case study. In *VISIGRAPP (5: VISAPP)*, pages 87–95.

Rodriguez, R., Dokladalova, E., and Dokladal, P. (2019). Rotation invariant cnn using scattering transform for image classification. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 654–658.

Rodriguez Salas, R., Dokladal, P., and Dokladalova, E. (2021). Rotation invariant networks for image classification for hpc and embedded systems. *Electronics*, 10(2):139.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Singh, C. and Singh, J. (2018). Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation. *Digital Signal Processing*, 78:376–392.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826.

Teague, M. R. (1980). Image analysis via the general theory of moments∗. *J. Opt. Soc. Am.*, 70:920–930.

Wang, Z., Yan, W., and Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE.

Xiao, B., Ma, J.-F., and Wang, X. (2010). Image analysis by bessel–fourier moments. *Pattern Recognition*, 43(8):2620–2629.

Xie, S., Girshick, R., Dollár, P., Tu, Z., and He, K. (2017). Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500.

Zhou, Y., Ye, Q., Qiu, Q., and Jiao, J. (2017). Oriented response networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.