International Science Group

ISG-KONF.COM

IV

INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE «TECHNOLOGIES, THEORIES AND DEVELOPMENTS: MODERN SCIENTIFIC TEACHING»

Valencia, Spain

September 23-26, 2025

ISBN 979-8-89814-218-6

DOI 10.46299/ISG.2025.2.4

TECHNOLOGIES, THEORIES AND DEVELOPMENTS: MODERN SCIENTIFIC TEACHING

Proceedings of the IV International Scientific and Practical Conference

Valencia, Spain
September 23-26, 2025

UDC 01.1

The 4th International scientific and practical conference “Technologies, theories and developments: modern scientific teachin” (September 23-26, 2025) Valencia, Spain. International Science Group. 2025. 183 p.

ISBN – 979-8-89814-218-6
DOI – 10.46299/ISG.2025.2.4

TABLE OF CONTENTS

REVIEW OF THE METHODS OF CLASSIFICATION OF YOGA EXERCISES IMPLEMENTED BY MACHINE LEARNING AND NEURAL NETWORKS

Podshyvalova Olha
Master in informatics
Kharkiv National University of Radio Electronics

Research supervisor:
Tvoroshenko Iryna
Ph.D., Associate Professor, Department of Informatics
Kharkiv National University of Radio Electronics

Classification methods based on machine learning are common in human pose classification tasks, since they have several advantages, among which are efficient work with large volumes of data of different types (in particular, photo and video data) and, on this basis, building models with high classification efficiency [1-4].

The Support Vector Machine (SVM) method is effective when applied to tasks that have a limited set of features. In the case of human motion classification, such features may be joint coordinates, angular values in joints, and motion trajectories.

The *k*-nearest neighbors (*k*-NN) method is also widely used in pose recognition and classification tasks [5-9]. Its essence lies in finding *k*-nearest neighbors for the classification object to determine which class it should be assigned to.

The advantage of decision trees and random forests is the determination of dependencies between the presented movements and types of asanas, which may not be linear [10].

The models presented above are effective at the stage of classifying a small number of static poses [11-15]. However, they will have problems with classifying asanas involving dynamic movements, since the machine learning approaches are limited when working with dynamic data.

The use of neural networks in visual information classification tasks is very common. A feature of the methods of this direction, which distinguishes them from machine learning methods, is the ability to determine key features from the input data stream independently. Such an approach makes the analysis process more flexible.

The choice of a specific type of neural network is determined by its specification. Next, methods such as convolutional neural networks, recurrent neural networks, hybrid models that combine several approaches into one method, and visual transformers will be presented.

Convolutional Neural Networks (CNNs) are one of the most widely used tools in visual material classification tasks, mainly for raster images. They make it possible to extract multi-level features, which increases classification efficiency. In the context of yoga pose recognition, CNN helps to effectively determine a specific asana from a single frame.

Another advantage of convolutional neural networks is the formation of a hierarchy of features. Such an approach makes it possible to correctly recognize an object even with changes in lighting, angle, zooming in or moving away from the camera, etc.

A disadvantage of this approach is that CNNs analyze only spatial structure and do not take into account the temporal sequence of movements, which limits their application in cases of dynamic object classification.

Recurrent architectures make it possible to analyze data sequences (for example, in the form of video data), taking temporal dependence into account.

Recurrent models, LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit), are applied for the classification of exercises that contain dynamic elements. Thus, such models will be effective in the classification of dynamic asanas, as well as in tracking transitions between different asanas, if we consider the use of the models discussed in this subsection in yoga pose classification tasks.

The use of RNN makes it possible to detect errors in technique during the movement itself, which opens prospects for creating interactive feedback systems. Their disadvantage is high computational complexity and sensitivity to long sequences. Although modern modifications of LSTM and GRU largely solve this problem.

The combination of convolutional and recurrent neural networks makes it possible to simultaneously consider spatial characteristics, that is, the type of static asana, and temporal dependencies, such as transitions between asanas or the performance of dynamic yoga exercises.

CNN is used for processing each frame of the video and extracting key features of the object, while at the same time, LSTM analyzes the dynamics of changes of these features over time. Such an approach makes the hybrid model capable of simultaneously combining the strengths of both methods.

Such models show especially high results in the recognition of yoga complexes or dynamic styles of yoga. However, a significant disadvantage of combining two types of models is the increased requirements for computational resources and the volume of training data, which complicates their training and further use.

Recently, transformers, originally created for natural language processing, have been successfully applied in computer vision as well.

Vision Transformers (ViT) split an image into small fragments, which are processed as sequences. This makes it possible to capture global spatial dependencies and patterns. Such an approach demonstrates high accuracy even in the classification of exercises that include non-standard positions of arms and legs. It should also be noted that transformers work well with large datasets, quickly adapting to the diversity of performance styles.

Another key advantage of the method under consideration is its ability to integrate multimodal data, which is, for example, combining video data and sensor data. This can provide an increase in the accuracy of evaluating the correctness of exercise performance in tasks similar to the one set in this work.

The main disadvantage of transformers, as well as of hybrid models, is the need for large computational resources and large training datasets. This complicates the use of this method in small-scale solutions.

References:

1. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set, *IEEE Access*, vol. 12, pp. 73376-73385.

2. Yakovleva O., Matúšová S., Tvoroshenko I., and Isaiev Y. (2024) Visitor counting based on video stream analysis from surveillance cameras to solve various business problems, *Verejná správa a regionálny rozvoj ekonómia, manažment a marketing*, XX(1), pp. 67-87.

3. Vaswani, A. et al. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

4. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, *Indonesian Journal of Electrical Engineering and Computer Science*, 33 (1), 113-125.

5. Гороховатський В., Передрій О., Творошенко І., Марков Т. (2023) Матриця відстаней для множини компонентів структурного опису як інструмент для створення класифікатора зображень, *Сучасні інформаційні системи*, 7(1), С. 5-13.

6. Pomazan, V., Tvoroshenko, I., and Gorokhovatskyi, V. (2023). Development of an application for recognizing emotions using convolutional neural networks, *International Journal of Academic Information Systems Research*, 7(7), pp. 25-36.

7. Gorokhovatskyi, V., Tvoroshenko, I., Kobylin, O., & Vlasenko, N. (2023). Search for visual objects by request in the form of a cluster representation for the structural image description, *Advances in Electrical and Electronic Engineering*, 21(1), pp. 19-27.

8. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2024) Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion, *Computers, Materials & Continua*, vol. 80, no. 2, pp. 3085-3106.

9. Gorokhovatskyi V., Tvoroshenko I. (2023) Identification of visual objects by the search request. *Decision making theory: proceedings of the international symposium, September 28, 2023*, Kyiv-Uzhorod, Ukraine, 25-27.

10. Gorokhovatskyi V., Chmutov Y., Tvoroshenko I., and Kobylin O. (2025) Reducing computational costs by compressing the structural description in image classification methods, *Advanced Information Systems*, vol. 9, no. 1, pp. 5-12.

11. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2025) Image description compression in classification structural methods, *IEEE Access*, vol. 13, pp. 43631-43641.

12. Tvoroshenko I., Gorokhovatskyi V., Kobylin O., and Tvoroshenko A. (2023) Application of deep learning methods for recognizing and classifying culinary dishes in images, *International Journal of Academic and Applied Research*, 7(9), pp. 57-70.

13. Gorokhovatskyi V., and Tvoroshenko I. (2024) An effective method for transforming an image description into a compact vector for classification. *Information*

Technology and Implementation (Satellite): Conference Proceedings, November 21, 2024, Kyiv, Ukraine, Publishing House «Caravela», pp. 25-28.

14. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2023) Statistical data analysis models for determining the relevance of structural image descriptions, *IEEE Access*, 11, 126938-126949.

15. Гороховатський В., Творошенко І., Сидоренко Д. (2021) Класифікація зображень із використанням кластерного подання, *Міжн. наук. симпозіум «Інтелектуальні рішення-С». Обчислювальний інтелект. Теорія прийняття рішень (Вересень 29, 2021)*. Київ – Ужгород, С. 44-45.