WSEAS TRANSACTIONS on ADVANCES in ENGINEERING EDUCATION
DOI: 10.37394/232010.2023.20.7

State-of-the-art CNN architectures for assessing Fine Motor Skills: A comparative study.

[1]KONSTANTINOS STRIKAS, [2]NIKOLAOS PAPAIOANNOU, [2]IOANNIS STAMATOPOULOS, [2]ATHANASIOS ANGEIOPLASTIS, [2]ALKIVIADIS TSIMPIRIS, [2]DIMITRIOS VARSAMIS, [1]PARASKEVI GIAGAZOGLOU

[1]Department of Physical Education and Sport Sciences
Aristotle University of Thessaloniki
Serres, GREECE

[2]Department of Computer, Informatics and Telecommunications Engineering
International Hellenic University
Serres, GREECE

Abstract: - It is considered that children's normal growth depends on their ability to use their fine motor skills. Deficits in fine motor skills in preschool children can interfere with even basic daily activities. Research also links these problems to future challenges. Therefore, early identification of preschool children's fine motoric abilities is considered essential. However, the assessment of the development of fine motor skills is considered to be a rather complex process. Complex and time-consuming methods are used for their reliable assessment, which also requires the presence of educational experts. The aim of this study is to investigate whether it is possible to create a simple and useful tool for assessing fine motor skills in preschool children, based on convolutional neural networks. For this purpose, a comparative study between 5 state-of-the-art CNN architectures is carried out, to investigate their accuracy in assessing fine motor skills. Drawings of Greek students from public kindergartens were used to train the investigated CNN models. The Griffiths II and the Eye Coordination Scale were used to assess the developmental age of preschool children. The findings demonstrate that, although challenging, automatic and precise detection of fine motor skills is feasible if a larger dataset is used to train deep learning models.

Key-Words: - Convolutional Neural Networks, Deep Learning, Fine Motor Skills, Preschool, Griffiths Test, Assessing Development

Received: June 15, 2022. Revised: February 16, 2023. Accepted: March 14, 2023. Published: April 12, 2023.

1 Introduction

Motor development refers to the change and enrichment of motor behaviour throughout life. The development of motor skills in children can be divided into two main categories: gross motor skills and fine motor skills. Gross motor skills refer to skills and movements that involve whole-body movements and large muscle groups to perform, while fine motor skills refer to precise movements that use smaller muscles, such as writing, drawing, tying shoelaces, etc. During the first few years of life, children grow rapidly and develop both gross and fine motor skills. Satisfactory development of fine motor skills is crucial for children's well-being. Several studies have shown that early difficulties in gross and fine motor skills have significant effects on academic performance and psychosocial maladjustment, [1]. They are also associated with reading development in primary school, [2], numeracy, [3], arithmetic, [4], mental imagery, [5], and working memory impairments, [6]. Various factors can prevent children from developing their fine motor skills to the full. One example is media use, which may have a negative impact on the development of fine motor skills in early childhood, [7]. In addition, obesity could also lead to lower fine motor precision performance, [8], while more generally the impact of modern society has also been studied as a factor that could potentially affect fine motor development, [9]. However, parental support for pre-school children, even if it is minimal, was considered to be really important to mitigate the consequences of fine motor impairments and to help children to reach their maximum potential in the development of fine motor skills, [10]. The same seems to be true for physical activity, which is considered to be beneficial for children, especially when provided regularly in a formal setting, [11], as well as for touch typing interventions, [12]. Therefore, the assessment of preschool children's fine motor skills is considered essential, considering the abovementioned effects on children's development. The Griffiths Scales No II is considered to be one of the most reliable developmental screening tests, [13]. It consists of six subscales to better assess motor development. In the bibliography, other reliable screening developmental tests such as Movement-ABC 2, [14], DSNR, [15], and pegboard tasks, [16], are also widely used. Recently, a new method using deep machine learning techniques, specifically convolutional neural networks (CNN), has been proposed for the assessment of fine motor skills in preschool children, called FineMotorSkillsCNN, [17]. Convolutional neural networks (CNN) are important tools for image recognition and classification tasks in various fields, such as smart agriculture, [18], medicine, [19], and self-driving cars, [20], probably because of their powerful image processing capability. Due to the stack of convolutional layers, CNNs can automatically extract features. Depending on the depth of the layers, CNNs could extract from low-level features, such as edges and dark spots, to high-level features, such as an object. Therefore, due to the success and high performance of CNNs in complex image processing problems, various methods have been proposed. In this paper, an evaluation of the performance of different state-of-the-art CNN architectures such as Efficient-Net, [21], ResNet, [22], VGG16, [23], and MobileNet, [24], for the assessment of fine motor skills of preschool children is presented, in order to investigate whether it could become a simple and useful tool for teachers and parents to detect possible impairments early and to help children reach their full potential in the development of fine motor skills. The remainder of the paper is organized as follows: Section 2, presents the basic elements of the theory and the methodology used, while Sections 3 and 4, present the results of the comparative study and the conclusions respectively.

2 Methods and materials

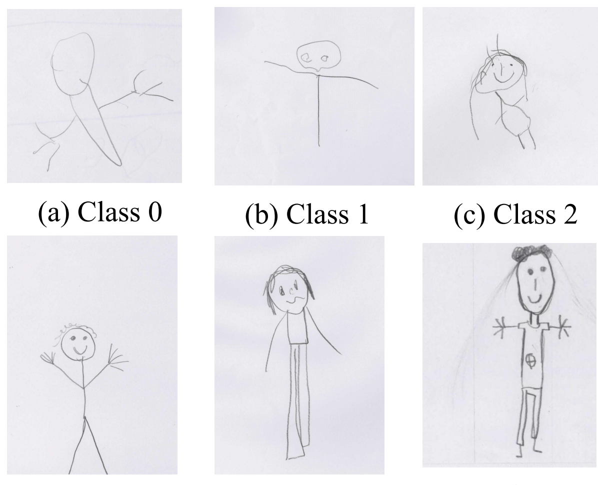
2.1 Dataset

The original dataset used in this study consists of 1601 images representing pictures of a man or a woman drawn by 442 children from 20 different preschool units. The dataset was divided into training and test sets. The training set, which was used to train the different CNN models, consists of 1121 images, while 480 images were used as the test set. As in [17] the images were divided into six different classes according to their developmental age, where class 0 to class 5 correspond to low and high developmental ages, respectively (Table 1).

Table 1: Griffiths II test scores and classes according to Developmental age (DA).

DA	Class
32-47	0
48-53	1
54-61	2
62-67	3
68-73	4
74-150	5

The classification of the pictures was carried out by three educational experts on the basis of the Griffiths II Scale D, which consists of six items for each year, such as threading beads onto a lace, building a tower of cubes, cutting with scissors, copying simple geometric shapes and drawing a house and a person freely. The developmental age (in months) was calculated by multiplying the number of items passed by two. The testing process started with simple tasks corresponding to a younger age and was stopped after six consecutive failures in six different skills. The developmental age of each child was then determined. Figure 1, shows pictures randomly selected from each of the six different developmental age classes.



(a) Class 0 (b) Class 1 (c) Class 2

(d) Class 3 (e) Class 4 (f) Class 5

Figure 1: Sample painting-class pairs from dataset.

2.2 State-of-the-art CNN architectures

The CNN architectures employed, are described in this chapter. The models studied were trained using a transfer learning approach.

2.2.1 VGG16

Attempting to improve on the original Convolutional Network Architecture proposed by Krizhevsky, [25], VGG addresses an important aspect of Convolutional Network Architecture, which is depth. VGG achieves high performance in both localisation and classification tasks by increasing the depth of the network (16 layers) and using small 3×3 convolutional kernels with convolutional stride 1 and 2×2 pooling layers with step size 2. Thus, the number of parameters increases significantly, while the decision function becomes more discriminative.

2.2.2 ResNet

ResNet primarily addresses the degradation problem. It has been found that as the depth of the networks increases, the accuracy decreases rapidly. To deal with this problem, ResNet introduced a deep residual learning framework that skips connections from initial filters to later layers. Identity mapping is performed on the shortcut connections. Based on the idea that if the model becomes deeper by adding layers constructed as identity mappings, the performance of the model shouldn't be worse than its shallower counterpart, ResNet prompts the network to approximate the residual function. Due to the small sample size of our dataset, ResNet50 is used in this study.

2.2.3 MobileNet

MobileNet is designed for mobile and embedded vision applications due to its low computational cost without compromising accuracy. To build lightweight deep neural networks, MobileNet differs from traditional approaches regarding convolution architecture. MobileNet is based on depth-separable convolutions, which split the process of both filtering and combining inputs into a new set of outputs into two distinct processes, a filtering layer and a combining layer. This is in contrast to standard convolution, which performs the entire process in one step. As a result, the parameters are drastically reduced and the model becomes less computationally intensive compared to a network of the same depth using regular convolutions.

2.2.4 EfficientNet

EfficientNet's baseline network is generated by a multi-objective neural architecture search, similar to MnasNet, [26]. Focusing on model scaling, it was found that performance can be improved by carefully balancing network depth, width and resolution. To achieve this, a compound scaling method was proposed, that scales network depth, width and resolution uniformly, using a set of predetermined constant scaling coefficients. Specifically:

$$
\begin{aligned}
d &= \alpha^{\phi} \\
w &= \beta^{\phi} \\
r &= \gamma^{\phi} \\
\alpha * \beta^2 * \gamma^2 &\approx 2 \\
&and \\
\alpha \geq 1, \beta \geq 1, \gamma &\geq 1
\end{aligned}
$$

where d, w, r correspond to depth, width and resolution respectively, α, β, γ are constants that can be identified using a simple grid search, while ϕ is a compound coefficient for uniformly scaling network depth, width and resolution. EfficientNet consists of 8 different models, namely EfficientNet-B0 to EfficientNet-B7, which are obtained by scaling the baseline mesh. In this study, EfficientNet-B0 and EfficientNet-B1 are used.

2.3 Transfer learning

Through the use of the machine learning task known as transfer learning, it is possible to transfer the knowledge gained from the training of a neural network in one problem to another task or domain, [27]. This method has been widely used to efficiently train models with limited data sets, and to overcome cost and time constraints. The use of pre-trained models instead of training neural networks from scratch speeds up the process, since the training model uses information from previous training processes and already understands the features of the problem under investigation. In summary, transfer learning produces more reliable and generalised models, while significantly reducing the risk of over-fitting.

2.4 Experimental setup

The study was conducted on a system running Windows 11 Pro, equipped with a 2nd Gen Intel(R) Core(TM) i7-12700 2.10GHz, 16GB RAM and a 1000GB SSD. Each of the models studied was trained for 15 epochs using the pre-trained model weights and an additional 5 epochs using its own weights to train on all three class scenarios. Adam was used as the optimiser, with a learning rate of 0.0001 and a loss function of categorical cross-entropy. Data augmentation techniques were used to extend the training set. Several techniques such as Horizontal RandomFlip, RandomRotation(0.2), RandomWidth(0.2), RandomHeight(0.2) and RandomZoom(0.2) were used to transform existing images. Tensorflow, Numpy, Keras, Matplotlib and sklearn libraries were used for the training and evaluation of the CNN models investigated in this study.

3 Results

As mentioned above, the main work of this study is to compare state-of-the-art CNN architectures for the assessment of preschoolers' fine motor skills. Due to the small sample size, and in order to further investigate the accuracy of the CNN models under study, the aforementioned classes of the dataset (see 2.1) were also merged into two and three classes. Thus, although it is assumed that the 6-classes case more accurately represents the groups of preschoolers studied, the 2-classes and 3-classes cases are also studied. In the case of two classes, the classes $0, 1, 2$ and $3, 4, 5$ from Table 1 were merged into classes 0 and 1 respectively. Similarly, in the case of three classes, classes $0, 1$ were merged into class 0, while classes $2, 3$ and $4, 5$ were merged into classes 1 and 2 respectively. The sample size for each case is shown in detail in Table 3. The training and testing accuracies of the investigated methods are presented both in tabular form (Table 2) and in graphical form (Fig. 2).

Table 2: Accuracy of the examined CNN architectures for a total of 2,3 and 6 classes.

Number of Classes	2		3		6	
	Training	Testing	Training	Testing	Training	Testing
EfficientNetB0	74.13%	68.94%	60.57%	38.12%	48.62%	22.92%
EfficientNetB1	75.74%	49.07%	63.07%	44.79%	46.74%	25.42%
MobileNet	70.21%	69.98%	48.53%	45.21%	30.95%	31.25%
ResNet	71.81%	55.69%	58.07%	53.33%	43.53%	30.00%
VGG16	70.21%	69.98%	50.76%	50.63%	29.88%	30.21%

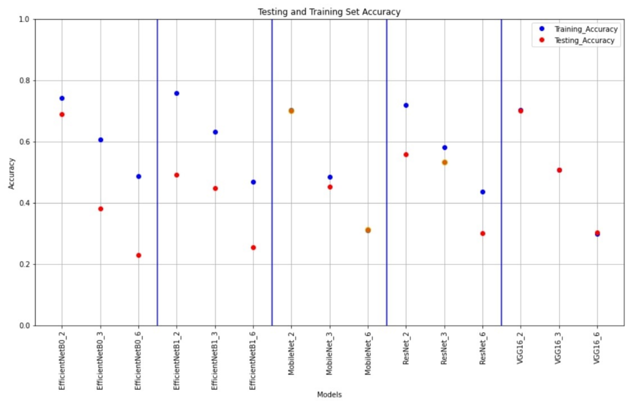


Figure 2: Accuracy of the examined CNN architectures, where `CNNmodel_x` refer to the number of classes.

The results suggest that as the number of classes increase, the accuracy of the examined methods decrease. For the case of two classes, two methods namely MobileNet and VGG16 both achieved the highest accuracy (69.98%) on testing data, while EfficientNet-B1 had the lowest accuracy (49.5%). Similarly to EfficientNet-B1, ResNet50 didn't performed well, with an accuracy equal to 55.69%. For the case of three classes, the accuracy of all the examined methods significantly decreased. In contrast to the 2-classes case, where ResNet50 performed poorly compared to the other methods, for the 3-classes case ResNet50 achieved the highest accuracy, namely 53.33%, while EfficientNet-B0 performed worse, with an accuracy equal to 38.12%, although it achieved the second highest accuracy on training data (60.57%). It should be noted that for the classification task of 3 classes, only ResNet50 and VGG16 achieved a performance of more than 50%. In the case of six classes, none of the examined methods managed to achieve an accuracy of more than 50%, neither on the test data nor on the training data. The highest accuracy was 31.25% achieved by MobileNet, while the second highest was 30.21% achieved by VGG16. Similar to the case of three classes, EffiecientNet-B0 performed poorly compared to the other investigated methods, achieving a classification accuracy of 22.92%.

Beyond the case of two classes, these results initially suggest a poor classification performance. This could be due to the small sample size. Table 3 shows that the number of images in the training and test sets is not evenly distributed. Specifically, out of a total of 483 images in the test set, only 12 correspond to class 0, while 30 and 49 correspond to classes 1 and 5 respectively. The above results for the 6-classes case indicate that ResNet50 achieved one of the highest performances, namely 30%. From its class distribution, which can be seen from the confusion matrix (Fig. 4: (b)), it can be seen that for class 3, out of a total number of 143 images, ResNet50 correctly classified 123 of them (86%). For classes 2 and 4, which also contained a large number of images in both the training and test sets, the model had a poor accuracy of 3% and 11.6% respectively. Nevertheless, for the second class, 86 out of 100 drawings were classified in the next class and for the fourth class, 111 out of 146 drawings were classified in the previous class. This doesn't seem to be the case for the classes mentioned above, to which a small number of drawings correspond (Table 3), because for class 0, the drawings weren't classified either in the correct class or in a neighbouring class. The same seems to be true for both the first and the fifth class. Therefore, this could be an indication that a larger data set is needed to increase the accuracy.

As already mentioned, in the case of six classes, even for classes 2 and 4, which contain a larger number of drawings compared to the other classes, ResNet50 had a poor accuracy of 3% and 11.6% respectively. For the same case, although EfficientNet-B1 had the second worst performance in general (25.42%), it is observed that for class 2 (see Fig. 4), out of a total number of 100 images, 34 were correctly classified (34%), while for class 4 the accuracy was around 15%. In addition, many of the images in the test set are classified into neighboring classes. Therefore, with the exception of class 3, for the classes with more images in the test set, EfficieNet-B1 shows a higher classification performance. It's worth noting that for class 1, EfficientNet-B1 was the only one to correctly classify one image and 17 out of 30 to the correct class or to a neighbouring class, in contrast to the other methods examined, for which, for the same case, they did not classify any image, either to the correct class or to a neighbouring class. Thus, despite its poor accuracy, EffiecientNet-B1 would tend to be considered more convincing than the other models examined, since it is considered to be more discriminative. The accuracy of VGG16, ResNet50 and MobileNet is significantly increased due to the fact that they classify most of the images of the test set in class 3, regardless of their actual class (see Fig. 4). This leads to a higher classification performance in general, since class 3 contains 143 out of 480 images of the test set and VGG16, ResNet50 and MobileNet achieve at least 85% accuracy in this class.

Figure 4 shows the rest of the confusion matrices of the investigated CNN models for the 6-classes case. Compared to MobileNet and VGG16, it can be observed that EfficientNet-B0 and EfficientNet-B1 are more discriminative. In particular, for class 2 case, MobileNet and VGG16 classify 91% and 97% of the drawings to the fourth class, while e.g. EfficientNet-B1 classifies 34% to the actual class and 1%, 40%, 10% and 15% to classes 1, 3, 4, 5 respectively.

The same seems to be true for all cases studied. For the case of 3-classes, it is observed that the models didn't correctly classify any of the images of class 1 (see Fig. 5 and Fig. 6). Table 2 illustrates that VGG16 had the second best accuracy, specifically 50.63%, but from Figure 5 it is observed that all the images are classified to class 2. Similar to VGG16, the MobileNet classifies most of the images to class 2 and the rest of them to class 3, achieving 45.21% accuracy. For the case of three classes, ResNet50 seems to perform better compared to the other methods, both in terms of accuracy and of the allocation of the images. ResNet50 achieved the highest accuracy, namely 53.33%. Figure 6 shows that ResNet50 classified correctly 0 out of 42 images of class 1, 155 out of 243 images of class 2 and 93 out of 195 images of class 3, contrary to VGG16, which achieved the second highest accuracy, but for classes 1 and 3 it classified correctly 0 out of 42 and 195 images respectively, while its performance improved since it classified all of the 242 images of class 2 correctly (see Fig. 5). Similar to ResNet, the EfficientNet-B0 and EfficientNet-B1 appeared to be more discriminative, since the images were allocated to different classes, while achieving 38.12% and 44.79% classification accuracy respectively.

Table 3: Sample size for each class of train and test set, where Train_x and Test_x, refer to train and test set respectively for the x-classes case.

Class	Train_6	Test_6	Train_3	Test_3	Train_2	Test_2
0	27	12	98	42	334	142
1	71	30	569	243	787	338
2	236	100	454	195	-	-
3	333	143	-	-	-	-
4	335	146	-	-	-	-
5	119	49	-	-	-	-

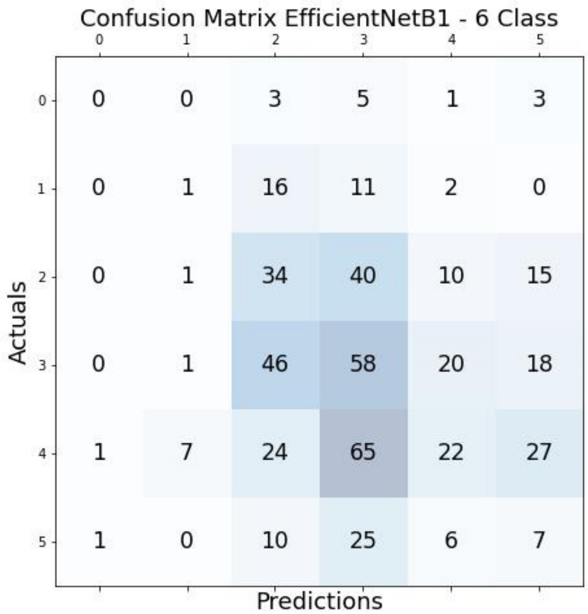
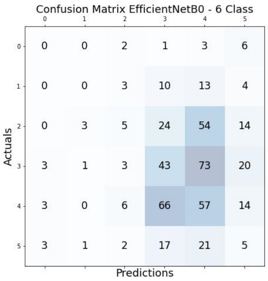
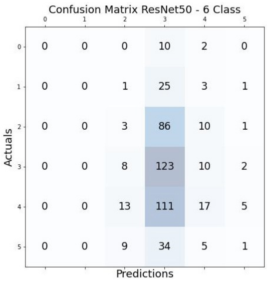


Figure 3: Confusion matrix of test set of EfficientNet-B1 for the 6-classes case. Columns and rows refer to predictive and actual values respectively.

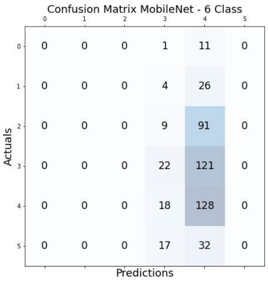
In summary, for the three cases studied, VGG16 achieved the highest average classification accuracy on testing data with 50.27%, followed by MobileNet with an average accuracy of 48.8%. On the testing data, the worst average classification accuracy was observed by EfficientNet-B1, which was limited to an accuracy of 39.76%. In contrast, on training data, EfficientNet-B1 outperformed the other CNN models examined with an average classification accuracy of 61.85%, followed by EfficientNet-B0 with 61.10%. In addition, MobileNet achieved a classification ac-



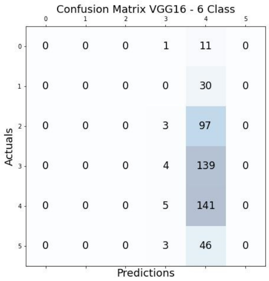
(a) EfficientNet-B0



(b) ResNet50

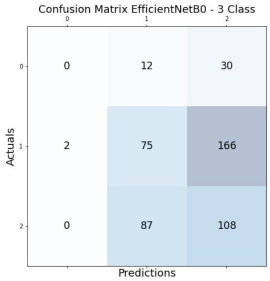


(c) MobileNet

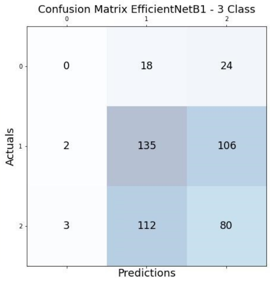


(d) VGG16

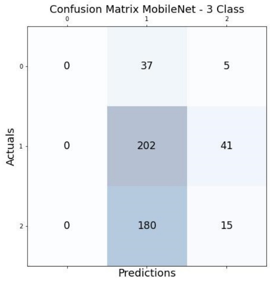
Figure 4: Confusion Matrices of the examined CNN architectures for the 6-classes case.



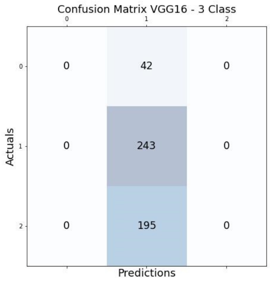
(a) EfficientNet-B0



(b) EfficientNet-B1



(c) MobileNet



(d) VGG16

Figure 5: Confusion Matrices of the examined CNN architectures for the 3-classes case.

curacy of 49.89%, which was the worst performance on training data.

4 Conclusion

This study compared five state-of-the-art CNN architectures to investigate their accuracy in assessing fine motor skills in preschool children. The data set used consists of drawings of children (male or female), classified by experts according to the children's developmental age. The Griffiths II and the Eye Coordination Scale were used to assess the preschool children's developmental age. As for the results, on testing data, besides that VGG16 achieved the highest average classification accuracy of 50.27% usually classifies most of data on a specific class. On training data, EfficientNet-B1 achieved the highest classification performance with an average classification accuracy of 61.85%. For the case of 3-classes ResNet50 seems to classify the drawings best. For the case of 6-classes EffiecientNet-B1 seems to perform better compared to the other models, since it allocates better the images to the correct classes or to the neighboring classes. In general, all the methods examined showed poor classification accuracy. This may be due to the small sample size. Confusion matrices confirm that the classes with larger sample sizes improve the classification, in contrast to the classes with fewer drawings in their image set. In summary, the results suggest that although challenging, automatic and accurate scoring of fine motor skills may be feasible when using a larger dataset.

5 Declarations

All authors declare that they have no conflicts of interest.

References:

[1] M. Katagiri, H. Ito, Y. Murayama, M. Hamada, S. Nakajima, N. Takayanagi, A. Uemiya, M. Myogan, A. Nakai, and M. Tsujii, "Fine and gross motor skills predict later psychosocial maladaptation and academic achievement," *Brain and Development*, vol. 43, no. 5, pp. 605–615, 2021.

[2] S. Suggate, E. Pufke, and H. Stoeger, "Children's fine motor skills in kindergarten predict reading in grade 1," *Early Childhood Research Quarterly*, vol. 47, pp. 248–258, 2019.

[3] U. Fischer, S. P. Suggate, and H. Stoeger, "Fine motor skills and finger gnosia contribute to preschool children's numerical competencies," *Acta Psychologica*, vol. 226, p. 103576, 2022.

[4] A. Asakawa and S. Sugimura, "Mediating process between fine motor skills, finger gnosis, and calculation abilities in preschool children," *Acta Psychologica*, vol. 231, p. 103771, 2022.

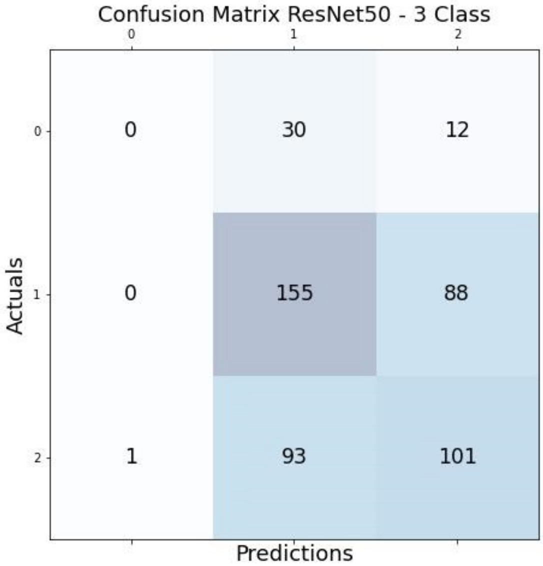


Figure 6: Confusion Matrix of ResNet50 architecture for the 3-classes case.

[5] P. Martzog and S. Suggate, "Fine motor skills and mental imagery: Is it all in the mind?," *Journal of Experimental Child Psychology*, vol. 186, pp. 59–72, 2019.

[6] E. Michel and S. Molitor, "Fine motor skill automatization and working memory in children with and without potential fine motor impairments: An explorative study," *Human Movement Science*, vol. 84, p. 102968, 2022.

[7] P. Martzog and S. Suggate, "Screen media are associated with fine motor skill development in preschool children," *Early Childhood Research Quarterly*, vol. 60, pp. 363–373, 2022.

[8] I. Gentier, E. D'Hondt, S. Shultz, B. Deforche, M. Augustijn, S. Hoorne, K. Verlaecke, I. De Bourdeaudhuij, and M. Lenoir, "Fine and gross motor skills differ between healthy-weight and obese children," *Research in Developmental Disabilities*, vol. 34, no. 11, pp. 4043–4051, 2013.

[9] D. Gaul and J. Issartel, "Fine motor skill proficiency in typically developing children: On or off the maturation track?," *Human Movement Science*, vol. 46, pp. 78–85, 2016.

[10] S. W. Bindman, L. E. Skibbe, A. H. Hindman, D. Aram, and F. J. Morrison, "Parental writing support and preschoolers' early literacy, language, and fine motor skills," *Early Childhood Research Quarterly*, vol. 29, no. 4, pp. 614–624, 2014.

[11] L. Dapp, V. Gashaj, and C. Roebers, "Physical activity and motor skills in children: A differentiated approach," *Psychology of Sport and Exercise*, vol. 54, p. 101916, 2021.

[12] H. McGlashan, C. Blanchard, N. Sycamore, R. Lee, B. French, and N. Holmes, "Improvement in children's fine motor skills following a computerized typing intervention," *Human Movement Science*, vol. 56, pp. 29–36, 2017.

[13] P. Giagazoglou, V. Tsimaras, E. Fotiadou, C. Evaggelinou, J. Tsikoulas, and N. Angelopoulou, "Standardization of the motor scales of the griffiths test ii on children aged 3–6 years in greece," *Child: care, health and development*, vol. 31, pp. 321–30, 06 2005.

[14] S. E. Henderson, D. Sugden, and A. Barnett, "Movement assessment battery for children-2," *Research in Developmental Disabilities*, 1992.

[15] H. Ito, W. Noda, S. Nakajima, Y. Tanaka, M. Hamada, and M. Katagiri, "Prediction of psychosocial maladaptation in elementary school based on development appraisal by nursery teacher: validation of the developmental scale for nursery record (dsnr)," *Japanese J Dev Psychol*, vol. 27, p. 59 – 71, 2016.

[16] J. Tiffin and E. Asher, "The purdue pegboard; norms and studies of reliability and validity.," *The Journal of applied psychology*, vol. 32, p. 234–247, 1948.

[17] K. Strikas, A. Valiakos, A. Tsimpiris, D. Varsamis, and P. Giagazoglou, "Deep learning techniques for fine motor skills assessment in preschool children.," *International Journal of Education and Learning Systems*, vol. 7, pp. 43–49, 2022.

[18] Y. Li, J. Nie, and X. Chao, "Do we really need deep cnn for plant diseases identification?," *Computers and Electronics in Agriculture*, vol. 178, p. 105803, 2020.

[19] J. Yang, L. Zhang, X. Tang, and M. Han, "Codnnet: A lightweight cnn architecture for detection of covid-19 infection," *Applied Soft Computing*, vol. 130, p. 109656, 2022.

[20] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array*, vol. 10, p. 100057, 2021.

[21] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," 2019.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," pp. 770–778, 06 2016.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, 09 2014.

[24] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," 04 2017.

[25] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Neural Information Processing Systems*, vol. 25, 01 2012.

[26] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. Le, "Mnasnet: Platform-aware neural architecture search for mobile," pp. 2815–2823, 06 2019.

[27] Y. Bengio, G. Guyon, V. Dror, G. Lemaire, D. Taylor, and D. Silver, "Deep learning of representations for unsupervised and transfer learning," vol. 7, 01 2011.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflicts of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.