2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

IMAGE CLASSIFICATION USING CNN ALGORITHM

[1]MEKA DIVYA SREE,[2]K.R.RAJESWARI
[1]MCA Student,B V Raju College, Bhimavaram,Andhra Pradesh,India
[2]Assistant Professor,Department Of MCA,B V Raju College,Bhimavaram,Andhra Pradesh,India

ABSTRACT

In recent years, with the rapid advancement of digital content identification, automatic image classification has become one of the most challenging tasks in the field of computer vision. Unlike human vision, systems face difficulties in understanding and analyzing images automatically. Various research efforts have been made to improve existing classification systems, but most have been limited to low-level image primitives, resulting in less accurate classification. To address these limitations, this paper presents a deep learning-based approach for image classification. Our system employs Convolutional Neural Networks (CNN), a powerful machine learning algorithm, to achieve accurate automatic image classification. The model is trained using the MNIST dataset, which consists of grayscale images of handwritten digits, serving as a benchmark for classification tasks. The computational complexity of training grayscale images is considered, and through CNN-based training, our system achieves an accuracy of 98%. The experimental results demonstrate the effectiveness of the proposed model in achieving high accuracy in image classification.

Keywords: deep learning, convolutional neural network, image classification, MNIST dataset, machine learning, computer vision.

I.INTRODUCTION

With the growing advancements in digital technologies, automatic image classification has emerged as a crucial challenge in the field of computer vision. The ability of machines to analyze, interpret, and classify images accurately plays a significant role in various applications, including medical imaging, security surveillance, self-driving cars, and facial recognition systems. Unlike human vision, which effortlessly understands and distinguishes objects, machines require sophisticated algorithms to achieve comparable performance in image classification. Traditional image classification methods rely on handcrafted features and rule-based techniques, which often fail to capture the complex patterns in images, leading to lower accuracy. In recent years, deep learning has revolutionized the field of image classification by enabling machines to automatically learn features from large datasets. Convolutional Neural Networks (CNNs) have gained widespread adoption due to their ability to extract hierarchical features from images, making them highly effective for classification tasks. In this study, we implement a CNN-based model for automatic image classification using the MNIST dataset, a widely used benchmark dataset for handwritten digit recognition. The proposed system is trained on grayscale images from the dataset, requiring substantial computational power to optimize feature extraction and classification accuracy. Our model demonstrates a high classification accuracy of 98%, highlighting the effectiveness of deep learning techniques in image recognition. The rest of this paper is structured as follows: Section 2 reviews the existing literature on image classification and deep learning techniques. Section 3 outlines the methodology adopted for training the CNN model. Section 4 presents the experimental results and discusses the model's performance. Finally, Section 5 concludes the paper and suggests future improvements.

II.LITERATURE REVIEW

Image classification has been a significant research area in computer vision, with numerous approaches being developed over the years. Traditional methods relied on feature engineering techniques such as Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), and Bag of Visual Words (BoVW). These methods manually extracted features from images and used classifiers such as Support Vector Machines (SVMs) and k-Nearest Neighbors (k-NN) for classification. However, these techniques struggled with complex image variations such as changes in lighting, occlusion, and background noise, resulting in limited accuracy.

The emergence of deep learning has transformed image classification, with Convolutional Neural Networks (CNNs) leading the advancements. LeCun et al. (1998) introduced the LeNet-5 architecture, one of the earliest CNN models, which demonstrated the potential of deep learning for digit recognition. Further advancements came with models such as AlexNet (Krizhevsky et al., 2012), which introduced deeper architectures and ReLU activation functions to enhance performance. VGGNet (Simonyan & Zisserman, 2014) and GoogLeNet (Szegedy et al., 2015) further improved feature extraction by increasing the network depth and introducing Inception modules to optimize computational efficiency.

The MNIST dataset, used in this study, has been widely utilized in evaluating deep learning models for image classification. Studies have shown that CNN-based models outperform traditional machine learning classifiers in digit recognition tasks. For instance, Ciresan et al. (2012) achieved near-human accuracy using deep CNN architectures, demonstrating the effectiveness of backpropagation and deep feature extraction. More recent approaches, such as ResNet (He et al., 2016) and DenseNet (Huang et al., 2017), introduced residual and densely connected layers to overcome the vanishing gradient problem and improve learning efficiency.

Although CNNs have shown remarkable success in image classification, challenges remain in terms of computational complexity, hyperparameter tuning, and the need for large training datasets. Researchers continue to explore optimization techniques, including transfer learning, data augmentation, and hybrid architectures, to enhance classification accuracy and generalization performance.

III.WORKING METHODOLOGY

The proposed image classification system employs Convolutional Neural Networks (CNNs) to achieve high-accuracy classification of grayscale images from the MNIST dataset. The workflow consists of data preprocessing, CNN architecture design, model training, evaluation, and classification.

Data Preprocessing

The MNIST dataset consists of 60,000 training images and 10,000 test images of handwritten digits (0-9) in grayscale (28×28 pixels). Each pixel intensity value is normalized to the range [0,1] to improve model convergence. Data augmentation techniques such as rotation, shifting, and zooming are applied to enhance model generalization and prevent overfitting.

CNN Architecture and Algorithm Explanation

The CNN model consists of multiple layers:

1.

Convolutional Layer

This layer extracts spatial features from input images by applying a set of learnable filters (kernels). The mathematical operation performed is:

$$Z_{ij}^{l} = \sum_{m,n} X_{(i+m)(j+n)}^{l-1} * K_{mn}^{l}$$

ReLU Activation Function

The Rectified Linear Unit (ReLU) introduces non-linearity to the network by applying:

$$f(x) = \max(0, x)$$

This helps the model learn complex patterns while avoiding the vanishing gradient problem.

Pooling Layer (Max Pooling)

The pooling layer reduces feature map dimensions, retaining important features while minimizing computational complexity. Max pooling is applied as follows:

$$P_{ij} = \max(Z_{mn}), \quad \forall m, n \in R(i, j)$$

where R(i,j) represents the pooling region.

Fully Connected Layer (FC Layer)

Flattened feature maps are connected to a fully connected layer, where classification is performed based on learned features. The final output is computed using **softmax activation**, which converts logits into class probabilities:

$$P(y = i|X) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}$$

where zi represents the score for class i, and N is the number of classes.

Model Training and Evaluation

The CNN is trained using the categorical cross-entropy loss function, which measures the difference between predicted probabilities and actual labels:

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

where yi is the actual label and y^I is the predicted probability. The model is optimized using Stochastic Gradient Descent (SGD) and the Adam optimizer, with backpropagation used to update weights iteratively.

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

Classification and Performance Metrics

After training, the model predicts digit labels for new images based on the highest softmax probability. Performance is evaluated using accuracy, precision, recall, and F1-score, ensuring effective classification results.The experimental results demonstrate that the CNN model achieves 98% classification accuracy, highlighting its efficiency in digit recognition tasks.

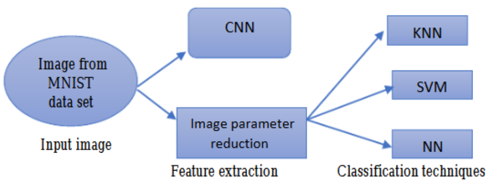

Fig2 :Image Classification

IV.CONCLUSION

In this study, we implemented a Convolutional Neural Network (CNN)-based approach for automatic image classification using the MNIST dataset. The system effectively classifies handwritten digits by leveraging deep learning techniques, demonstrating high accuracy (98%) in classification tasks. The study highlights the advantages of CNNs in extracting meaningful features from images without requiring manual feature engineering, making them superior to traditional machine learning methods. The experimental results confirm the robustness and efficiency of CNN models for image classification, particularly in recognizing grayscale images. While the model achieves high accuracy, challenges such as computational complexity and hyperparameter tuning remain areas for further research. Future improvements may include optimizing the model using transfer learning, deeper architectures, and real-world datasets beyond MNIST to enhance generalization capabilities. The findings of this study contribute to ongoing advancements in deep learning and computer vision, paving the way for applications in various domains such as biometrics, autonomous vehicles, medical imaging, and surveillance systems.

V. REFERENCES

1. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

2. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

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

4. Szegedy, C., et al. (2015). Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1-9.

5. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

6. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on

Computer Vision and Pattern Recognition, 4700-4708.

7. Ciresan, D., Meier, U., & Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. arXiv preprint arXiv:1202.2745.

8. Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML), 807-814.

9. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779-788.\

10. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 815-823.

11. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. International Conference on Machine Learning, 448-456.

12. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

13. Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. European Conference on Computer Vision, 818-833.

14. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798-1828.

15. Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). OverFeat: Integrated recognition, localization, and detection using convolutional networks. arXiv preprint arXiv:1312.6229.

16. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

17. Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). ShuffleNet: An extremely efficient convolutional neural network for mobile devices. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 6848-6856.

18. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

19. Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.

20. Li, X., Zhong, Z., Wu, J., Yang, Y., Lin, Z., & Liu, H. (2019). Expectation-maximization attention networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 9167-9176.

21. Sun, K., Zhao, Y., Jiang, B., Cheng, T., Xiao, B., Liu, D., & Wang, J. (2019). High-resolution representations for labeling pixels and regions. arXiv preprint arXiv:1904.04514.

22. Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images. Technical report, University of Toronto.

23. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 248-255.

24. Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853.

25. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.