Hand written Digit Recognition Using Machine Learning Techniques

التعرف على الأرقام المكتوبة بخط اليد باستخدام تقنيات التعلم الآلي

Yasir Madhi Hasan

ياسر ماضي حسن

yasir.madhi٧٩@gmail.com

Mustansiriyah University. Iraq

أشراف / الدكتور خالد صافي

AUL Arts, Sciences & Technology University in Lebanon

Khaled.safi@innovai-tek.com

المستخلص

يعد التعرف على الأرقام المكتوبة بخط اليد مهمة بالغة الأهمية وصعبة في مجال الرؤية الحاسوبية. وله تطبيقات واسعة النطاق في مجالات مثل معالجة المستندات، وإدخال البيانات آليا، وإنشاء نماذج التعلم الآلي. تستكشف هذه الأطروحة تعقيدات التعرف على الأرقام باستخدام الشبكة العصبية التلافيفية (CNN) لفك رموز الأنماط المعقدة التي ترى في الأرقام المكتوبة بخط اليد. وتقدم دراسة شاملة للأدبيات تحليلا مفصلا للتطور التاريخي والوضع الحالي لأساليب التعرف على الأرقام. ويقدم البحث تعريفا دقيقا للمشكلة، مبرزا أهمية هذا العمل في سياق معالجة المعلومات الرقمية وتطبيقات الذكاء الاصطناعي.

تتمثل الأهداف الرئيسية في إنشاء نموذج CNN متقدم، مع التركيز الدقيق على إعداد البيانات، وطرق التوسيع، وهيكل النموذج. وتدعم فعالية النموذج اختبارات وتحليلات شاملة، تكشف عن مستويات رائعة من الدقة ومقاييس الأداء. ولا تعزز هذه الأطروحة النقاش الأكاديمي حول التعرف على الأرقام فحسب، بل لها أيضا تداعيات عملية على التطبيقات العملية. يتناول هذا البحث الإنجازات المتحققة، مسلطا الضوء على التطور المستمر لتقنيات التعرف على الأرقام. ويحدد العمل المستقبلي، مع التركيز على مجالات مثل تحسين أداء النماذج، وتحقيق قابلية التفسير من خلال الذكاء الاصطناعي القابل للتفسير، والتعرف على العديد من اللغات والنصوص، ومواكبة أحدث التوجهات التكنولوجية. وتقدم هذه الأطروحة بحثا شاملا في التعرف على الأرقام المكتوبة بخط اليد، مما يساهم في التقدم المستمر والاستفادة من الذكاء الاصطناعي في مجالات متعددة.

الكلمات المفتاحية: التعلم الآلي، الكتابة اليدوية، التعرف على الأرقام، الشبكة العصبية التلافيفية.

Abstract

Handwritten digit identification is a significant and challenging task in the field of computer vision. It has wide-ranging applications in areas like as document processing, automated data entry, and the creation of machine learning models. This thesis explores the complexities of digit recognition by utilizing a Convolutional Neural Network (CNN) to decipher the intricate patterns seen in handwritten numbers. A thorough literature study provides a detailed analysis of the historical development and current state of digit recognition methods. The research presents a strong issue definition, highlighting the importance of this work in the context of processing digital information and applications of artificial intelligence.

The primary goals guide the creation of an advanced CNN model, with careful focus on data preparation, augmentation methods, and model structure. The model's effectiveness is supported by

thorough testing and analysis, which reveal impressive levels of accuracy and performance metrics. The thesis not only enhances the scholarly discussion on digit recognition but also has practical ramifications for real-world applications. The conclusion examines the accomplished milestones, highlighting the ongoing development of digit recognition technologies. The future work is outlined, focusing on areas such as improving model performance, achieving interpretability through explainable artificial intelligence, recognizing several languages and scripts, and aligning with upcoming technology trends. This thesis provides a thorough investigation into the recognition of handwritten digits, which contributes to the continuous progress and utilization of artificial intelligence in several fields.

Keywords: Machine Learning, Hand written, digit Recognition, Convolutional Neural Network.

Background and Motivation

Both machine learning and deep learning have been very helpful to the areas of AI and computer science. Machine learning and deep learning make detection, learning, and prediction easier and faster. In order to recognise handwritten digits (٩-٠) from the famous MNIST dataset, this research evaluates and contrasts several classifiers and their parameters. These include KNN, PSVM, NN, and convolution neural network.

Computer scientists are exploring deep learning and machine learning in an attempt to build more intelligent computers. Repetition is the key to human learning; we can achieve anything if we do it enough. Then all of a sudden, his neurons in the brain start firing in sync, and the operation is over in no time. Additionally, it is quite similar to deep learning.

Various neural network topologies are used for different issues. Applications include image and audio segmentation, object detection, item identification, and many more.

Thanks to advancements in handwritten digit recognition, computers are now able to read digits written by humans. This is a difficult task for the computer to do since handwritten numbers are imprecise and varied. This problem may be solved by using handwritten digit recognition, which can identify a digit from a picture. A trained digit recognition system might have practical uses in many areas, such as online handwriting recognition on computer tablets or systems, recognising number plates, and handling hand-filled forms.

Problem Statement

The purpose of this thesis is to apply Convolution Neural Network ideas toward the development of a model that can extract the digits from a handwritten image. The first focus is on building a model that can identify numbers, but this may easily be expanded to encompass letters and even a person's handwriting. The suggested system's primary focus is on learning how to apply machine learning to the problem of handwriting recognition.

Research Objectives

In order to overcome the shortcomings of current methods, this study will compare and contrast several machine learning techniques for handwritten multidigit string recognition. In order to provide a more comprehensive and unbiased comparison, assessment factors such as accuracy, f ١score, and confusion matrix are used. This study not only compares machine learning's performance on classification tasks, but it also examines the effects of various data organisation strategies and data formats on model performance, laying the groundwork for future studies that aim to train end-to-end frameworks optimally in such scenarios .

Literature Review Introduction

Handwriting recognition technologies are increasingly finding many uses. Letters, bank cheques, and other handwritten documents can't be scanned and preserved without handwriting recognition technology. Another helpful and often used tool is online handwriting recognition software .

Tablets and other electronic devices may greatly enhance educational programmes that use handwriting recognition. Specialised software has been made possible by handwriting recognition technology for individuals with cognitive or motor disabilities, whether they are children or adults. Some of the most used applications on mobile phones include handwriting recognition algorithms. Instantaneously digitise any nearby handwritten texts with the use of your phone's camera. You may search or translate the digital writings into several languages.

There is an amazing breadth to text recognition. The subfields of handwriting recognition and typewriter/computer writing recognition are both part of the larger area of handwriting recognition. A computer or typewriter's recognising field is faster and more accurate. Number and letter recognition are more likely to achieve high success rates than handwriting recognition since these data types do not include distinguishing lines and patterns.

Cursive optical character recognition (OCR) is a subfield of optical character recognition (OCR) that focuses on digitising handwritten text and instructions. In addition, "handwritten digit recognition" refers to a computer's ability to identify and classify human-written numerals from a variety of inputs (such as images, documents, and touch displays) into one of nine distinct number groups. Pattern matching has long served as the basis for identifying handwritten numbers. The first programme to utilise pattern recognition in this manner was built by Moscow resident Shelia Guberman in ١٩٦٢.

Machine Learning Overview

One popular technique for extracting text from images and labels is optical character recognition (OCR). You won't need deep learning to solve this computer vision challenge, unlike a few others [١]. Many optical character recognition systems were therefore in use prior to the ٢٠١٢surge in popularity of deep learning. Modern optical character recognition (OCR) software utilises artificial intelligence (AI) and machine learning to accomplish even higher levels of accuracy in tasks such as language identification, reading handwritten text and writing styles, traversing typical data restrictions, and many more. It's reasonable to say that optical character recognition has been greatly affected by deep learning. Recent years have seen tremendous progress in the field, with Convolutional Neural Networks (CNNs) becoming a powerful tool for solving computer vision problems in many different contexts.[٢]

Problems with handwritten digit recognition include: - Users' handwriting is often inconsistent and changes over time.

Cursive handwriting makes it hard to separate and recognise individual letters.

While printed language is legible in any direction, handwritten writing is susceptible to a variety of twists and turns.

-Inadequate source picture quality might lead to inaccurate recognition results. This is the reason why many machine learning algorithms have problems with number recognition.

Types of machine learning :

Supervised Learning

Consider the fascinating realm of machine learning, where the intricate dance between input and output data unfolds. At its core, the process involves feeding an algorithm an input dataset, a collection of information that serves as the foundation for the system's learning. This dataset essentially becomes the fuel for the algorithmic engine, guiding it through the intricate landscape of patterns and relationships.

The magic takes place whilst the set of rules begins to decipher those styles, drawing connections and making educated guesses approximately the corresponding output facts. In the world of supervised getting to know, a particular facet of this fascinating domain, the connection between input and output is express and strong. Each piece of enter data is meticulously paired with its corresponding end result, forming a symbiotic dating that the algorithm exploits to refine its predictive prowess. This symbiosis implies that the set of rules, given a fixed of inputs, cannot best determine styles but additionally make correct predictions approximately the related outputs .

Unsupervised Learning

Delve into thefascinating realm of unsupervised learning, a paradigm Where in algorithms embark on a journey of discovery within the big expanse of enter facts. Unlikesupervised mastering, right here, the algorithm faces the fascinating assignment of decoding latent structures and patterns with out the crutch of categorized or categorized enter statistics. Picture an algorithm as anintrepid explorer navigating uncharted territories, armed most effective with the raw, unannotated data at its disposal.In this autonomous quest for know-how, the set of rules engages in a technique of self-discovery, unraveling the inherent intricacies and hidden relationships within the statistics. The absence of predefined labels liberates the algorithm, permitting it to operate with a experience of independence and adaptableness. It becomes a virtual detective, tirelessly sifting via the facts, figuring out similarities, and grouping elements based totally on their intrinsic traits.

Reinforcement Learning

Step into the dynamic realm of reinforcement mastering, where the primary objective is to unravel the intricacies of behavior to figure the actions that maximize rewards within a given scenario. Unlike different machine mastering paradigms, reinforcement learning adopts a awesome technique by way of focusing on the concept of trial and blunders, a great deal similar to how human beings analyze via enjoy. The algorithm, similar to an adaptive agent, navigates via an surroundings, making choices and taking moves based totally at the cutting- edge situation. The remaining quest is to decipher the most effective direction of motion to undertake under specific situations, a task that involves a non-stop technique of gaining knowledge of and edition .

Deep Learning

"Deep Learning is a superpower," remarks Arthur Andrew Ng. It has a wide range of potential
applications, including computer vision, artistic creation, language translation, medical diagnosis, and the construction of autonomous driving car components. "That is, without a doubt, a superpower.[١١] "
Deep learning employs a variety of machine learning techniques focused on learning representations of data rather than algorithms designed for particular tasks [١١]. Supervised and unsupervised learning are the two main categories [١٢, ١٣]. It's a set of methods used in machine learning to create many data representations, one for each possible degree of abstraction. A multi-layer approach may be used to calculate complex, nonlinear functions .

The data is passed on from one layer to another by means of calculation and transformation as it is received as input by each layer. There is a wide variety of connections between neurons in various levels and across layers of the same network .[١٤ ,٦]

The foundation of deep learning is the idea of mimicking the brain's operations in order to develop better learning algorithms, which will bring about a new age of advancements in AI and machine learning [٧]. Deep learning's popularity has been on the rise due to recent technological advancements.

Support Vector Machine (SVM)

The first description of the Support Vector Machine (SVM) was given by Boser, Guyon, and Vapnik in ١٩٩٢[١٥]. It is a crucial component of several ML methods. Support vector machines (SVMs) are linear classifiers used in supervised learning that may be used to binary classification as well [١٠]. So, SVM isn't only for machine learning; it's also used extensively in areas like natural language processing, picture analysis, CV, and voice recognition. By creating a perfect hyperplane, it enhances the degree to which the two classes in the data are separated [١٦]. An excerpt from the training data used to verify optimal decision-making positions [١٧]. This is the stage when regression and classification prediction tools really shine, as overfitting might happen if the algorithm isn't watched closely. Applying the strategy effectively to training data allows us to verify its efficacy using test data [١٧]. It is mostly because of [١٨] that SVM outperforms other algorithms. Support vector machines (SVMs) have recently shifted their attention from classification to regression, while they were originally designed to handle classification issues.[١٩]

Artificial Neural Network (ANN)

Conceptualize the artificial neural network as a tricky mesh intricately woven from interconnected cells, each akin to a computational unit designed to system and transmit information .[٢٠]

This network shape mirrors the complicated internet of neurons in the human mind, wherein each cellular, or neuron, plays a important function in data processing. As facts flows through the community, every cell receives inputs and transmits them to the next mobile within the collection, creating a continuous comments loop that characterizes the community's dynamic learning technique. Assigned an activation feature (denoted as "A"), each cellular strategies incoming information, introducing non-linearities and permitting the community to capture complicated styles and relationships within the information [٢٠]. To visualize the architecture and capability of a mature synthetic neural community, you'll be able to talk over with a block diagram that showcases the interconnectivity of neurons and the drift of facts within the network. Notably, these neurons are not isolated entities however are systematically linked in a sequential fashion, forming layers that collectively contribute to the community's capacity to realize and extract significant insights from data [٢٠]. The network, in its superior shape, carries multiple hidden layers [٦], strategically scaled to suit the granularity of inputs while juxtaposed with the dataset. This layering scheme enhances the network's capability to determine hierarchical capabilities and tricky styles, making it adept at dealing with complicated datasets and extracting treasured facts for a myriad of applications.

Convolutional Neural Network (CNN)

A class of multi-layer neural networks known as the Convolutional Neural Network [٢١] become created to improve the processing of two-dimensional information, together with snap shots and movies. Their primary determinant is the distribution in their weights throughout time, and they are on the whole designed to deal with time collection and speech [٢٢]. In order to teach correctly, there are numerous levels of hierarchy. This approach leverages spatial and temporal correlations to lessen the quantity of parameters that ought to be learnt, making it an development upon normal feed-forward backpropagation schooling. Its stated motive as a deep getting to know device is to want minimum information instruction .[٢٢]

The first layer of a CNN's hierarchical layout takes as inputs visual characteristics that have been retrieved domestically. The "hidden layers" that link the network's enter and output nodes are known as "transition nodes," and the intermediate nodes are referred to as "hidden layers." The hidden layer methods the inputs, at the same time as the output layer determines the magnificence chances. As the quantity of input or the number of layers in a standard neural network increases, it becomes capable of doing more complicated computations. These comprehensive parameter estimates were designed to prevent overfitting results .[٢٢]

But CNN's interlayer connections aren't perfect. Only a small region could receive messages from the neurons in this layer. By stacking several layers, we may increase the degree of abstraction from the most basic to the most complex levels, while also fostering the local spatial link in the data.[١٨]

While lower-level layers can only see a portion of the input data, higher-level levels may access the whole dataset and draw conclusions from it. There are three different types of layers used by convolutional neural networks (CNNs):

.١Convolutional layer

.٢Pooling layer

.٣Fully-connected layer

Other Machine Learning Algorithms

Random Forest

The random woodland method is beneficial for solving regression and class issues, similar to selection timber. Building and using several selection timber to get average results is the way to move in line with paintings logic. Because it introduces a random factor into the tree-constructing procedure, this method receives its call from that characteristic. It compares a randomly chosen subset of traits in opposition to the complete set if you want to choose the pleasant characteristic to appoint when splitting a node. This causes a greater variety of tree kinds to emerge.[٢٣]

Decision Tree

Enter the realm of problem- solving with the versatile tool referred to as choice bushes, where the intricacies of classification are unraveled through the advent of a hierarchical version. These trees serve as a scientific roadmap, breaking down complicated troubles into a sequence of easy, sequential decisions. The beauty of decision trees lies of their clarity and simplicity, making them an attractive choice for a numerous range of problem- fixing eventualities.

K-Nearest Neighbor

The K-Nearest Neighbour (KNN) approach is a versatile tool that may lead you on a journey into the realm of classification by using the values of recent observations in a sample set that contains recognised classes. It is an outstanding decision-maker for new observations.

Essentially, KNN is based on the proximity principle, which evaluates how much the training data ،which consists of previously assigned classes to observations، and the new test data point، which is still awaiting classification، overlap.

Related Work

Using metrics like accuracy, complexity, execution time, and epoch count, the authors of [٢٧] examine and contrast several neural network classifiers for handwritten digit recognition. The MNIST dataset served as the foundation for their study, which included using some fundamental picture preprocessing methods like as scaling and filtering. But they've discovered that adding more preprocessing steps to their study yields far superior results. Results show that the test programme is generally accurate (٪١٠٠-٩٨), with industrial photographs ranging from ٪٩٨-٩٦accurate. All of the suggested algorithms were within one percent of one another in terms of accuracy. The most accurate, but computationally expensive, option is the Convolutional Neural Network (CNN). In [٢٨], the authors examine and contrast several approaches to training and evaluating models that can recognise the numbers ٠٣٩when written by hand. Among these techniques are support vector
machines (SVMs), Bayesian networks, Random forests, and Multi-layer perceptrons (MLPs) .

The MLP classifiers were unable to proceed due to overfitting and local minima. Even though it took longer to train the model with additional data, they discovered that the accuracy increased. Additionally, when comparing two classes, they discovered that SVM offered the most substantial margins of separation. Although LeNet ٤, a CNN technique, had the quickest detection time (٠,٠٥ms), it required a relatively lengthy training period (five weeks) and had the highest accuracy.(٪٩٩,٣)

The CNN method using the Rectified Linear Units (ReLU) activation function is evaluated by the authors of [٢٩] using the MNIST dataset, which contains handwritten numbers. They use the same DL٤J environment as we use. Two convolutional layers yielded the best results; the first layer had ٣٢filters and a ٥٥-by-٧٧-pixel window size, while the second layer had ٦٤filters and a ٧٧-by-٧٧-pixel window size. The end result is a model with a ٪٩٩,٢١accuracy rate. The authors of this research dare to say that their approach produces stellar results in every category.

CNN Architecture Components

Since their introduction a decade ago, Convolutional Neural Networks have revolutionized several areas of pattern identification, including image processing and speech recognition. One of CNNs' main benefits is that they allow an ANN to rely less on a large number of parameters. This achievement has inspired researchers and developers to take on increasingly complicated and large-scale models, which standard ANNs were previously unable to handle. When attempting to resolve difficulties, CNN relies heavily on the fact that the attributes being employed are not location-specific. To restate, it is not required to pinpoint the precise location of the faces in the photos for a face detection program to work. If they are in the pictures, then discovering them is all that matters. When input is delivered down via the layers of a CNN, abstract features are retrieved. One method for classifying images involves looking for edges first, then moving on to more complicated forms, and lastly to higher-level characteristics like faces, as shown in Figure ٧and discussed in.[٤٠-٣٨]

Convolution

The network is currently being fed raw picture data. In the hidden layer of a Multi-Layer perceptron, for instance, a link between the input layer and a single neuron requires ٣٢*٣٢* ٣weight connections for the CIFAR- ١٠dataset.

Pooling Fully

The concept of pooling, which seeks to streamline future layers, relies heavily on down- sampling to achieve its goals. For those familiar with image processing, this is akin to reducing the resolution.

Connected

A classic neural network's neuronal connections are analogous to the fully-connected layer .

All of the nodes in a fully connected layer are linked to each other and to all of the nodes in the layer below and above it.

Evaluation Metrics

In the expansive landscape of machine learning, the iterative process of testing and refining models stands as a cornerstone for achieving optimal performance. The efficacy of a model is intricately tied to its ability to tackle the specific challenges presented by the problem at hand. Consequently, an array of measures is employed to meticulously evaluate the model's effectiveness and fine-tune its predictive capabilities.

Accuracy

Classification accuracy stands as the cornerstone in evaluating the effectiveness of a model tasked with categorizing data. At its essence, this fundamental parameter encapsulates the model's ability to make accurate predictions by quantifying the proportion of instances that are correctly labeled.

Precision

Precision, a pivotal metric in classification assessment, elucidates the proportion of accurately identified positive instances by our model in relation to all instances labeled as positive. In essence, it quantifies the model's precision in correctly identifying instances of the positive class.

F-Score

In classification evaluation, the F-score (also called the F١-score) provides a balanced estimate of the recall and accuracy of a model.

This harmonic mean of precision and recall provides a holistic understanding of a model's performance.

Dataset Description

The data files train.csv and test.csv.The CSV file has grayscale pictures of hand-drawn digits ranging from zero to nine.

Each picture has dimensions of ٢٨pixels in height and ٢٨pixels in width, resulting in a total of ٧٨٤pixels. Every pixel is assigned a single pixel-value, which represents the brightness level of the pixel. Higher pixel-values correspond to darker pixels. The pixel-value is an integer ranging from ٠to ٢٥٥, including both endpoints.

The training data collection, named train.csv, contains a total of ٧٨٥columns. The initial column, referred to as the "label", represents the numerical value that was manually selected by the user. The remaining columns consist of the pixel values of the corresponding picture.

Each column in the training set is labeled with a name in the format "pixelx", where x is an integer ranging from ٠to ٧٨٣, including both endpoints. In order to determine the position of this pixel inside the picture, let us assume that we have broken down the x-coordinate into the form x = i * ٢٨+ j, where i and j are whole numbers ranging from ٠to ٢٧, including both endpoints. Pixelx is positioned at the intersection of row i and column j in a ٢٨x ٢٨matrix, using zero-based indexing.

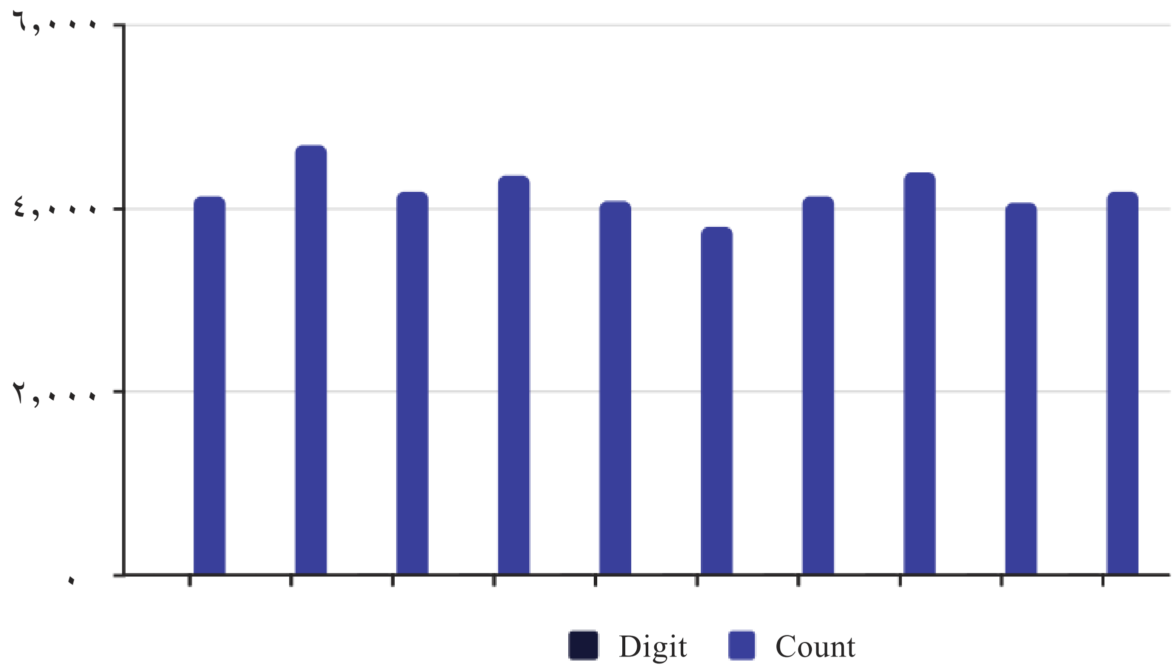
For instance, the pixel labeled as pixel ٣١corresponds to the pixel located in the fourth column from the left and the second row from the top, as seen in the ascii-diagram below.

The test data set, denoted as test.csv, is identical to the training set, with the exception that it lacks the "label" column.

Exploratory Data Analysis

Let us do exploratory data analysis in order to gain a deeper understanding of the dataset. In figure ٢٠ we present the number of data in train and test data.

In figure ٢١ we present the number of each features in data set. The graph above displays a consistent distribution of classes for the MNIST dataset.



Data Preprocessing and CNN Model Construction

The first phase is extracting the features (X) and labels (y) from the training dataset. Significantly, the color depth of the photos is altered from ٠ to ١, thereby transforming the grayscale images into a binary format. Normalization is performed in this procedure to rescale the pixel values within the range of ٠ to ١. This helps improve the convergence of the model during training. In addition, the photos undergo a conversion from grayscale to RGB format, which is a crucial step for models that require input in color. The X array obtained is subsequently transformed into a four-dimensional structure to fit the picture dimensions.

Moreover, the algorithm offers a glimpse into the range of pixel values included in the training dataset, displaying both the lowest and highest values. The

knowledge of pixel intensity distribution and alignment of the normalization procedure with the dataset's features heavily relies on these crucial parameters .
The last example plot provides a visual depiction of a training image associated with a specific label, offering a concrete illustration of the preprocessed data. This extensive preparation procedure guarantees that the training dataset is properly structured and standardized, establishing the foundation for resilient model training.
The model is constructed in a sequential manner, with layers being added one by one. The first layers
include of two convolutional layers followed by a max-pooling layer, which contribute to the extraction of features and reduction of dimensionality. The chosen activation function is Rectified Linear Unit (ReLU), which is renowned for its efficacy in bringing non-linearity to the model. In addition, dropout layers are incorporated after the max-pooling layers with a dropout rate of ٠,٢٥. This is done to mitigate overfitting by randomly deactivating a portion of the neurons throughout the training process.
The next layers replicate this pattern by incorporating a further set of convolutional layers, max-pooling, and dropout, hence improving the model's capacity to capture hierarchical characteristics within the data. Subsequently, the flattened layer readies the data for insertion into a highly linked neural network. The model has two densely connected layers with Rectified Linear Unit (ReLU) activation and dropout to enhance feature extraction and regularization. The model's output is generated by the last dense layer, which utilizes softmax activation to provide class probabilities .

Results, Conclusion and Future Work

The model is compiled using the Adam optimizer, a popular choice for gradient-based optimization algorithms. The loss function selected is categorical crossentropy, appropriate for multiclass classification tasks. Additionally, accuracy is chosen as the evaluation metric to assess the model's performance during training. The model.compile statement brings together these components, preparing the CNN for the subsequent training phase.

The model.fit function is then employed to train the CNN on the provided training data (x and y) for a specified number of epochs (٢٠ in this case) with a batch size of ٦٤. The validation data (x_test andy_test) are also incorporated to evaluate the model's performance on unseen data during training .

The training process is monitored, and the training history is stored in the variable history. Subsequently, the code employs Matplotlib to visualize the training and validation loss over epochs in one plot and the training and validation accuracy over epochs in another plot. These visualizations offer insights into the model's convergence, potential overfitting, and generalization performance. The accuracy for this model is ٠,٩٩ and the loss is .٠,٠٠٢٩

Conclusion

The crowning glory of this thesis on handwritten digit identification indicates a noteworthy success inside the area of pc vision and sample popularity. By thoroughly inspecting several tactics and techniques, we have efficiently deciphered the intricacies concerned in identifying handwritten numerals. The thorough literature analysis established the premise via imparting treasured insights into the historic development and modern breakthroughs in digit popularity. The trouble description clarified the importance of this work, specially in the context of numerous applications which include virtual document processing and gadget getting to know model introduction.

Future Work

The effective fruits of this thesis on the recognition of handwritten digits establishes the inspiration for numerous interesting strains of future research and investigation. First and predominant, the endeavor to improve version performance remains of extreme significance. Refining the existing Convolutional Neural Network (CNN) structure or investigating more state-of-the-art neural community topologies, consisting of recurrent neural networks (RNNs) or attention tactics, may additionally bring about additional upgrades in accuracy and generalization.

Moreover, the use of switch gaining knowledge of techniques, utilising pre-educated models on huge datasets, provides a promising opportunity. Utilizing statistics derived from models trained on various datasets has the potential to improve the version's capacity to become aware of complex patterns in handwritten numbers. Utilizing ensemble getting to know techniques, which include merging predictions from one of a kind fashions, can beautify the resilience and dependability of the consequences.

References

[١]B. V. Dasarathy, "Nearest neighbor (NN) norms: NN pattern classification techniques," IEEE Computer Society Tutorial, .١٩٩١

[٢]E. Hancer, B. Xue, and M. Zhang, "Differential evolution for filter feature selection based on information theory and feature ranking," Knowledge-Based Systems, vol. ١٤٠, pp. ١١٩-١٠٣, .٢٠١٨

[٣]Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. ٥٢١, no. ٧٥٥٣, p. ٤٣٦, .٢٠١٥

[٤]Y. Bengio, "Learning deep architectures for AI," Found. Trends® Mach. Learn., vol. ٢, no. ١, pp. ١–١٢٧, .٢٠٠٩

[٥]R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. ٢٢, no. ١, pp. ٦٣–٨٤, .٢٠٠٠

[٦]J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. ٦١, pp. ٨٥–١١٧, .٢٠١٥

[٧]Bengio, Yoshua; LeCun, Yann; Hinton, Geoffrey (٢٠١٥). "Deep Learning". Nature. ٥٢١(٧٥٥٣): ٤٣٦–.٤٤٤

[٨]Bengio, Y.; Courville, A.; Vincent, P. (٢٠١٣). "Representation Learning: A Review and New

Perspectives". IEEE Transactions on Pattern Analysis and Machine Intelligence. ٣٥(٨): ١٧٩٨ ١٨٢٨.

[٩] C. J. Burges and B. Schölkopf, "Improving the accuracy and speed of support vector machines,"in Advances in neural information processing systems, ١٩٩٧, pp. ٣٧٥–٣٨١.

[١٠]V. Vapnik, Estimation of dependences based on empirical data. Springer Science & Business Media, ٢٠٠٦.

[١١]M. I. Fanany, "Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM)," in Information and Communication Technology (ICoIC٧), ٥ ٢٠١٧th International Conference on, ٢٠١٧, pp. ١–٦.

[١٢]A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. Lang, "Phoneme recognition using timedelay neural networks," IEEE Trans. Acoust., Speech, Signal Processing, vol. ٣٧, pp. ٣٢٨–٣٣٩, ١٩٨٩.

[١٣]Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R. Wu, A.Y., "An optimal algorithm for approximate nearest neighbor searching in fixed dimension," Journal of the ACM, ٢٠١٧.

[١٤]Kirmizigul Caliskan, S., Sogukpinar, I., KxKNN : "K-ortalama ve K-en yakin komsu yontemleri ile aglarda nufuz tespiti," Electronic Journal of Vocational Colleges (Ejovoc), Kırklareli , ٢٠١٥.

[١٥]O. M. Khanday and S. Dadvandipour, "Analysis of machine learning algorithms for character recognition: a case study on handwritten digit recognition," Indonesian Journal of Electrical Engineering and Computer Science, vol. ٢١, no. ١, pp. ٥٨١-٥٧٤, ٢٠٢١.

[١٦]S. Ali, Z. Shaukat, M. Azeem, Z. Sakhawat, and T. Mahmood, "An efficient and improved scheme for handwritten digit recognition based on convolutional neural network," SN Applied Sciences, vol. ١, no. ٩, pp. ٩-١, ٢٠١٩.

[١٧]O. Voloshchenko and M. Plechawska-Wójcik, "Comparison of classical machine learning algorithms in the task of handwritten digits classification," Journal of Computer Sciences Institute, vol. ٢١, ٢٠٢١.

[١٨]S. Ghosh, "Comparative Analysis of Boosting Algorithms Over MNIST Handwritten Digit Dataset," in Evolutionary Computing and Mobile Sustainable Networks: Springer, ٢٠٢٢, pp. ٩٩٥-٩٨٥.

[١٩]H.-h. Zhao and H. Liu, "Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition," Granular Computing, vol. ٥, no. ٣, pp. ٤١٨-٤١١, ٢٠٢٠.

[٢٠]K. Senthil Kumar, S. Kumar, and A. Tiwari, "Realtime Handwritten Digit Recognition Using Keras Sequential Model and Pygame," in Proceedings of ٦th International Conference on Recent Trends in Computing, ٢٠٢١: Springer, pp. .٢٦٠-٢٥١

[٢١]M. B. Abdulrazzaq and J. N. Saeed, "A comparison of three classification algorithms for handwritten digit recognition," in ٢٠١٩International Conference on Advanced Science and Engineering (ICOASE), ٢٠١٩: IEEE, pp. .٦٣-٥٨

[٢٢]R. KARAKAYA and S. KAZAN, "Handwritten digit recognition using machine learning," Sakarya University Journal of Science, vol. ٢٥, no. ١, pp. ٧١-٦٥, .٢٠٢١