Overview of Algorithms for Image Recognition

Abstract

The significance of image recognition technology is highlighted by its wide applications in fields such as security, medical image analysis, and data analysis. Its growing popularity reflects advancements in research. Traditional machine learning methods have markedly improved feature extraction, while deep learning techniques have advanced significantly due to the application of various neural networks. This paper reviews algorithms and systems for image recognition, covering both traditional and deep learning methods. It provides extensive descriptions of classification and object detection techniques involving feature extraction, convolutional neural network designs, and neuron activation functions. The focus extends to traditional algorithms like k-nearest neighbor, support vector machine, Naive Bayes, and parallel cascade selection. Additionally, it explores various deep learning approaches for image interpretation, detailing different convolutional network dimensions and neuron model constructions. The paper concludes by illustrating algorithms with application examples and clarifying the differences between traditional methods and deep learning.

Keywords: Image Recognition, Deep Learning, Object Detection, Feature Extraction, Convolutional Neural Networks.

1. Introduction

Image recognition is the intriguing and complex process of automatically identifying and recognizing various objects present within digital images and video content. This rapidly growing field is gaining immense importance due to the enormous potential it holds for a large number of practical applications across a variety of industries, as well as tremendous market opportunities that are emerging. The continually expanding volume of visual content generated daily, alongside the increasing requirements for efficient processing in diverse domains, drives the creation of sophisticated and efficient algorithms capable of analyzing and processing raw visual data with remarkable speed and accuracy[1]. In contrast, traditional image processing techniques and computer vision algorithms typically require years of iterative experimentation to establish and achieve good performance outcomes. This lengthy process results in high costs and makes these traditional approaches less suitable for direct deployment in real-world applications. In sharp contrast, current and modern image recognition systems, particularly those that utilize deep learning methodologies, have made significant advancements in multiple fields. They provide cutting-edge tools that are not only highly efficient but also exceptionally accurate in terms of performance, bringing forth new possibilities and capabilities in the realm of automated object recognition. [2]

The design of efficient, robust, and highly versatile algorithms specifically tailored for recognizing a wide variety of objects present within digital images and videos constitutes an exceedingly intricate and multifaceted undertaking. This inherent complexity arises from several diverse factors, including but not limited to the substantial variability in the quality and characteristics of the images that are available for processing, the array of diverse computational resources that may or may not be accessible in various contexts, as well as the pressing necessity for real-time performance in a vast and expanding array of practical applications [3]. In order to effectively address these significant challenges, substantial research efforts have culminated in the definition and refinement of a series of increasingly effective, well-optimized techniques and methodologies. However, it is important to note that until the last decade unfolded, a considerable amount of experimental work was essential to successfully develop algorithms that could operate efficiently at performance levels that transcended mere laboratory-use paradigms, thus paving the way for more widespread and practical applications in real-world scenarios[4]. The recent advent of sophisticated image recognition algorithms, especially those pioneered and developed based on deep learning methodologies, has made it possible to achieve dramatically improved performances in this highly specialized field. These advancements continue to transform how we approach image and video analysis, presenting new opportunities and challenges to researchers and practitioners alike.[5]

Over the years, the methods used in Image recognition have completely changed from conventional feature-based methods to deep learning algorithms. SIFT and HOG for instance relied on feature extraction to train models such as SVM and Random Forests. But CNN which is implemented by Alex Net ResNet shifted the paradigm by introducing end to end learning and achieving great results on big datasets. Currently, Vision transformers have extensively improved the effectiveness and performance metrics by making use of self-attention mechanisms. Today, image recognition is used in a variety of fields such as autonomous vehicles and medical imaging, and makes progress in solving issues related to the need for massive amounts of data; expensive computing, and model explain ability and all of this leads to better and more ethical and accessible innovations

2. Fundamentals of Image Recognition

In this section, we cover the basics of image recognition for those without a technical background. Image recognition follows a step-by-step process. The initial step, image preprocessing, enhances the image quality by highlighting key areas and minimizing noise. Feature extraction, another critical step, identifies and quantifies essential information from the image. This reduces the large image data size to a low-dimensional feature vector, which is crucial for efficient recognition. Once preprocessing and feature extraction are complete, various algorithms can carry out pattern recognition [6]. While deep learning algorithms are central to modern image recognition, the aforementioned processes are vital for a broad range of recognition algorithms. During preprocessing, noise—an unwanted disturbance affecting the signal—needs to be minimized. Techniques to reduce noise include the Wiener filter, Gaussian filter, and median filter. Another technique is resizing the image; smaller images require less time and resources for subsequent recognition. Edges in an image represent significant intensity changes and mark the start and end of objects. Feature extraction involves identifying important information in the image by analyzing pixel light intensity values. Key image properties such as shape, outline, corners, texture, orientation, and position can be derived from these intensities. [7]

2.1. Image Preprocessing

In image recognition, increased dimensional and topological features can hamper performance, necessitating preprocessing to enhance efficiency. This includes improving presentation and ensuring clear output. Key steps involve grayscale conversion, normalization, resizing, and noise reduction. Normalization aligns pixel grayscale levels, reducing contrast and mitigating spatial illumination variations, thereby enhancing average intensity. Image size is critical; sizes too far from the target can decrease recognition ability. Resizing standardizes spatial dimensions, simplifying variable input sizes. Colour spaces may be reduced, retaining minimal data relative to the original image, with techniques chosen to minimize content loss for subsequent processing stages. However, excessive resizing can lead to information loss. [8]

Image noise reduction cleanses inputs from camera sensor noise and lighting gaps. Pixel colors may be filtered, but averaging can blur details. A combination of averaging and Gaussian filters mitigates the downsides of both. This study focuses on optimal filters for grey reduction, as noise affects recognition accuracy. Unique inputs are essential for pure images to achieve better results. Techniques like erosion and dilation significantly diminish noise, while averaging and Gaussian filters help reduce image noise. However, excessive narrowing might result in detail loss, affecting system recognition. [9]

Balancing noise reduction with resolution helps maintain recognition efficiency while optimizing hardware resource use. Intensive preprocessing can lower computing time in image recognition, significantly altering quality and average output. The interplay between preprocessing techniques and image feature extraction is fundamental for effective image analysis and recognition system design [10]. Image size impacts the recognition system's ability to discern small details, necessitating adaptable preprocessing techniques in recognition systems. The range of pixel values directly reflects the reproducible detail level, defining preprocessing methods aligned with various recognition goals and timeframes.[11]

2.2. Feature Extraction

Feature Extraction Timing: Feature extraction is the process of identifying and quantifying characteristic information that distinguishes differences between the objects or scenes. It is the basis of the image analysis process[12]. Many popular feature extraction methods for image recognition include edge detection, histogram analysis, watershed analysis, wavelet transform, and so on. Some descriptors such as SIFT, HOG, and LBP have been introduced to image recognition and are popular due to their robustness and high accuracy. [13]

The importance of feature dimensions in the classification problem for the different classes of images can be analysed. Here, manifolds separating the different classes of samples from one another are highlighted with different colons [14]. This indicates the role of the eigenvectors and their corresponding eigenvalues in the classification performance of the algorithms. It is without a doubt that in a number of cases, the features that mostly affect the classification performance are the additional features, and these features exploit further any variations between the known feature classes in the form of the underlying manifold [15]. These additional features might be quite useful to build a classifier with a higher discriminating power. Feature extraction is important because images have high volume, variability, and complexity. Identifying characteristics that distinguish an object or scene of interest is key to achieving recognition that is both accurate and less sensitive to image variability such as changes in views, lighting, etc. [16]

Feature extraction in (figure 1) is the first step in image recognition to render the set of image data into the associated feature space. There are two main dominating methods of feature extraction, namely traditional methods and machine learning methods. Traditional methods such as edge detection, thresholding, and region of interest extraction are simple low-level techniques to get image regions that might be relevant to the system [17]. On the other hand, machine learning methods such as deep learning and clustering have been proposed to produce high-level image features with agreement on the complete data distribution in the image. These methods do not depend on mathematical assumptions.Only when domain knowledge is available to distinguish important traits can features be extracted. [18]

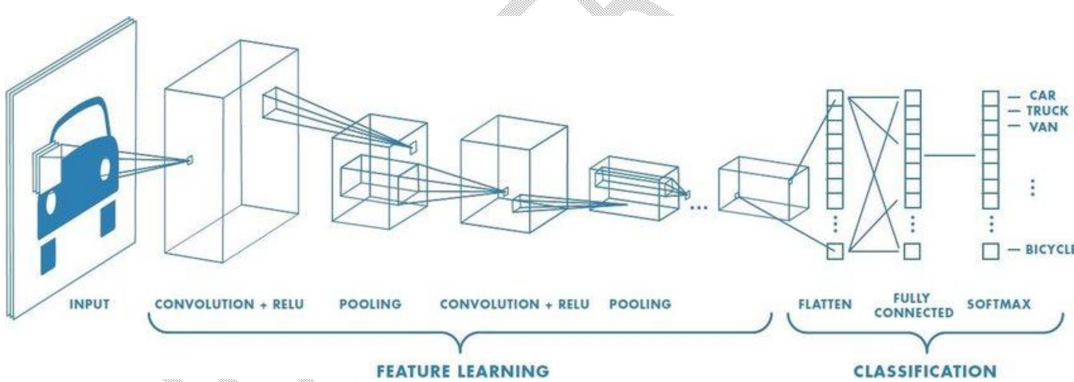

Figure 1: Layer of Feature Extraction

3. TraditionalImage Recognition Algorithms

Traditional image recognition algorithms are those that were in use before the development of deep learning techniques. Several well-established algorithms were used in this domain and are the focus of a survey in this section. Details regarding the method and the evaluation have been extracted in this survey. Earlier algorithms were shallow algorithms, and traditional benchmarks consisted of the evaluation of algorithm performance against a wide set of binary classifiers [19]. Together with an additional set of descriptive and performance metrics (sensitivity, specificity, accuracy, F1-Score, Receiver Operating Characteristic, Area Under the Curve, confusion matrix, precision-recall curves, false positive rate, false negative rate, and other descriptive statistics), this binary evaluation allowed us to develop a richer view of algorithm performance on binary classification tasks. [20]

Several traditional algorithms were employed for binary classification, notably SVM, which supports non-linear kernels beneficial for behavioural analysis, fraud detection, and characterizing skin in images. Comparisons between K-NN and SVM indicated that K-NN generally enhanced classification performance, though it faced challenges with increasing computational complexity as dataset sizes grew. K-NN's efficacy was particularly noted in an OCR algorithm evaluation, outperforming SVM and Naive Bayes. It excels when prior distributions are unknown, feature-to-sample ratios are high, and no dominant class exists [21]. Moreover, K-NN offers flexibility in implementation and classification across numerous classes. However, it is sensitive to irrelevant features and performs poorly when dissimilar classes cluster closely in feature space, making it less suitable for tasks with few classes or significant irrelevant features. Traditional methods provided a foundational framework for modern algorithms, primarily addressing image classification and basic object detection with limited depth. They laid the groundwork for advanced algorithms and deep learning techniques by paving the way through empirical and methodological details. This iterative improvement fostered the development of related methods. Empirical studies also utilized the ROC curve, which assists in visualizing classifier performance and serves as a comparative tool in ROC space, illustrating performance gradients effectively(2) . [22]

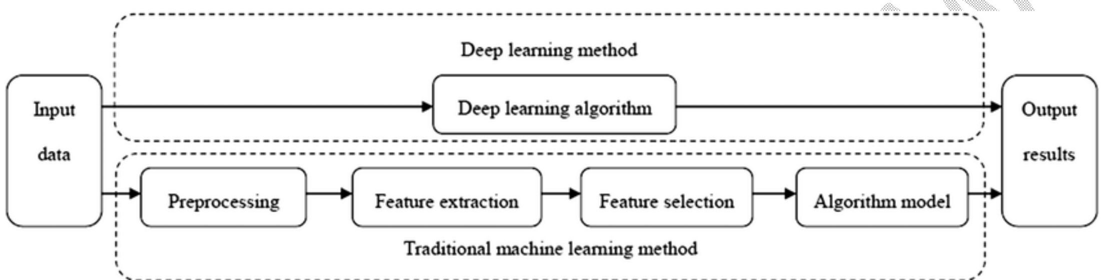

Figure 2:Deep learning and conventional machine learning are evaluated.

3.1. Support Vector Machines

Support vector machines (SVM) are effective algorithms in image recognition, focusing on finding the optimal hyperplane, the SVM classifier, for classifying linear and nonlinear data through a geometric approach. The SVM classifier maximizes the margin between two classes to minimize misclassification. Key advantages of SVM include reducing overfitting, effectively handling higher-dimensional data, and binary classification applications, and making it useful across various research areas, including plant and animal identification and medical image processing. SVM's effectiveness was noted in tuning raw data domain features [23]. However, its computational demands can complicate processing large image databases, and it is sensitive to noise due to its design focusing on fewer noisy data points. Other conventional algorithms, like k-NN and k-means, show less capability in comparison, as SVM can analyses high-dimensional features from image databases. Despite its challenges, kernel SVMs are used for feature determination and optimizing parameters through spectral analysis feature rejection [25]. (Figure 3 & 4)supports vector machines over random forest models for classifying remote sensing images. [24]

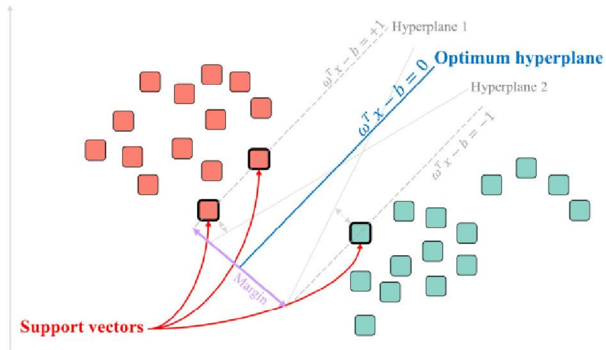

Figure 3: SVM data sample that is linearly separable.

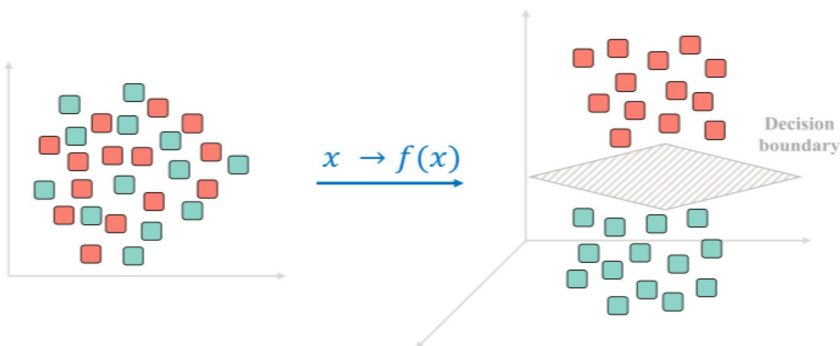

Figure 4: SVMis an illustration of data that can be separated nonlinearly using the kernel method .

3.2. K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a key algorithm in image recognition, classifying query data points using their 'k' nearest neighbors. Its effectiveness in high-feature spaces has made KNN popular for small datasets and prevalent in recommendation systems, search engines, semantic analysis, and associative memory. Variants include weighted neighbor classifiers and KD-trees, with KNN excelling in multi-class classification tasks. A variety of distance measures can compute distances between examples. However, KNN faces limitations like choosing k-nearest neighbours, distance matrices, and the curse of dimensionality [26]. Optimization methods, such as feature selection and ranking, can mitigate these challenges. KNN has proven useful in recognition problems, showing robustness to k choices, including handwritten digit and facial recognition applications.KNN's fundamental tenet is "like with like," meaning that the query point's nearest neighbors usually fall into the same class. There are multiple processes in the classification process, as demonstrated in (Figure 5,6). [27]

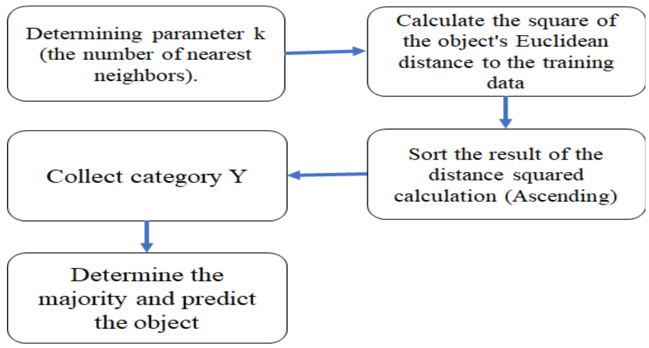

Figure 5: K-NNSteps in Analyzing

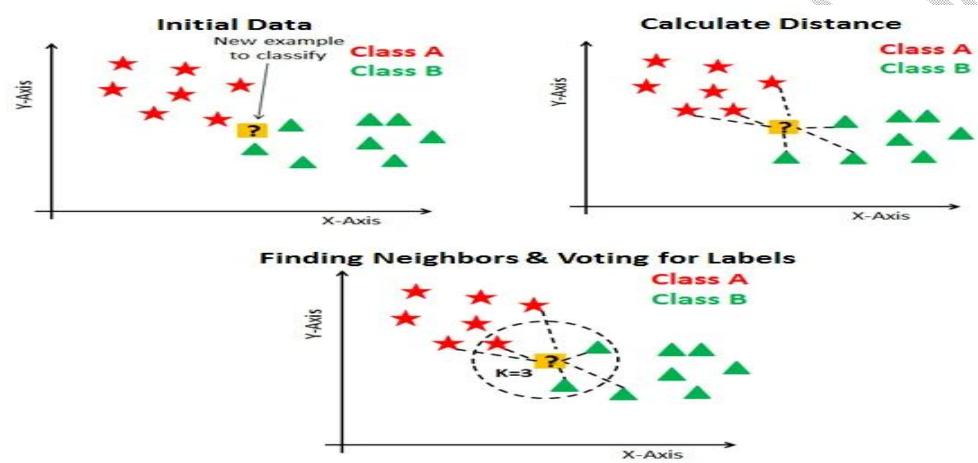

Figure 6:A Comprehensive Guide to the KNN Algorithm: An Understanding of Its Foundations and Uses.

4. Deep Learning-Based Image Recognition

Image research has entered a golden age with many new algorithms emerging annually, supported by robust libraries and datasets for training and evaluation. A clear trend is the rise of deep learning algorithms, dominating image recognition techniques through efficient architectures and hardware [28]. Convolutional Neural Networks (CNNs) have transformed the field, excelling in feature learning tasks. Early systems relied on manual feature extraction but struggled with large datasets, leading to a demand for fully automated feature extraction. Neural networks largely addressed these challenges. The journey of image recognition began in the 1960s, hampered by high costs and poor-quality data, but saw significant improvements in the 1980s and 1990s due to advancements in hardware and camera technology [29]. By the early 2000s, statistical approaches yielded successful recognition systems. The evolution of image recognition has paralleled advancements in informatics and AI. Traditional pattern recognition methods based on expert knowledge were gradually replaced bydescriptor-based learning, where distinguishing features are automatically learned and applied [30]. Displayed in (Figure 7) provides the general structure of the CNN-BILSTM network for speech and picture emotion identification.

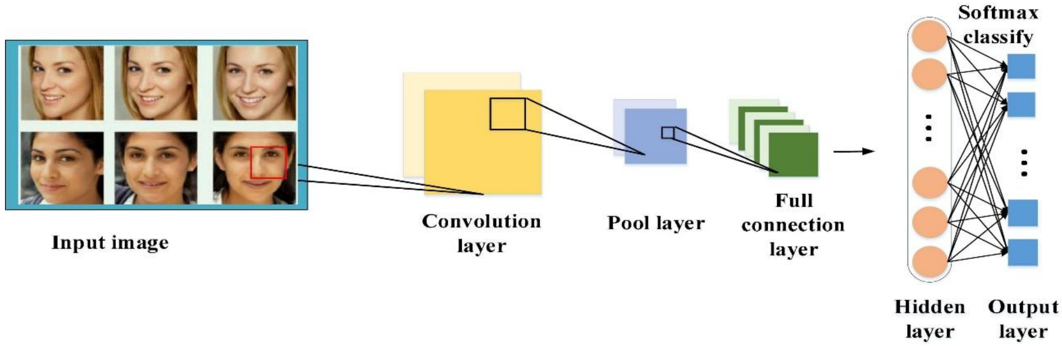

Figure 7: Image Recognition-Based Deep Learning.

.

4.1. Convolutional Neural Networks

In image recognition, convolutional neural networks (CNNs) have been one of the most transformative architectures. CNN models can automatically and adaptively learn spatial hierarchies of features from images by learning a variety of sophisticated filters. This led to their design, which consists of multi-layer neural networks, in which convolutional layers, subsampling layers, and fully connected layers were arranged in a chain with an end-to-end component. Convolutional layers mainly consist of convolutional operations by applying convolutional filters to the input feature matrix, followed by the addition of a bias term and an activation function, which has usually been the Rectified Linear Unit [31]. Deep neural networks, such as CNNs, have become so successful largely due to a series of innovations in the use of gradient-based optimization techniques. This differs greatly from the difficulties or impossibilities that hand-engineered features face for complicated tasks. Today, CNNs have an advantage in many applications in computer vision and image recognition, such as image classification, object detection, image segmentation, and other recognition tasks based on data patterns or structures. For CNN image recognition algorithms, either a pre-trained model can be deployed directly or re-training can be applied using a fine-tuning or transfer learning strategy [32]. In (Figure 8), a Convolutional Neural Network (CNN) is a deep learning algorithm that processes input images, learns features through trainable weights and biases, and distinguishes objects. Unlike traditional methods requiring manual feature engineering, CNNs automatically learn features with minimal pre-processing.

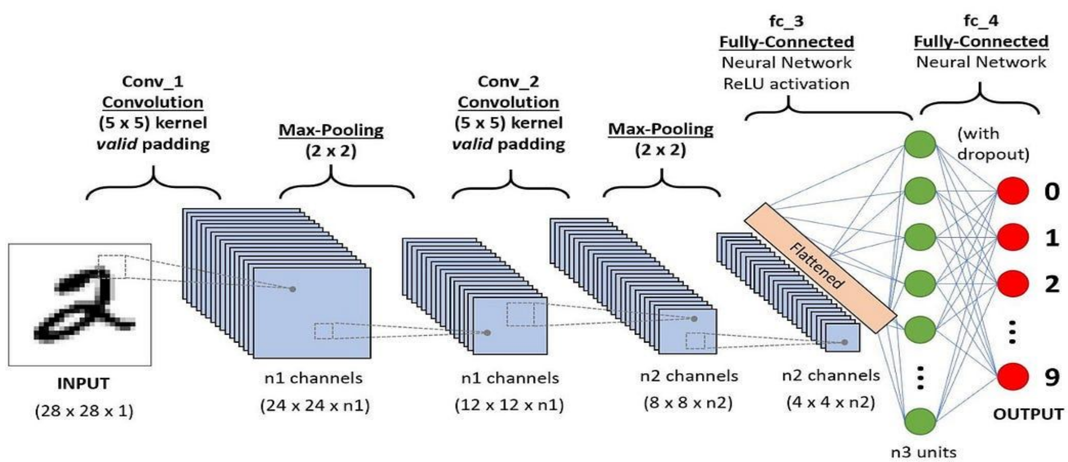

Figure 8:CNN & Convolutional Neural Networks are Used for Image Recognition.

In convolutional neural networks, the main challenges have been in their training, including the following: (1) The learning process of a deep network has been quite time-consuming, particularly in the case of a huge scale of data to handle. (2) For deep models, large-scale data with higher computational costs to collect and process has been urgently needed. Convolutional neural networks are seeing massive uptake in applications including security, medical organizations, gaming, internet services, and autonomous vehicles [33]. In the healthcare domain, CNN models have been deployed to detect diseases of the retina from fundus image data. Another application of CNN in autonomous vehicles is deploying the model to accurately classify other road users and pedestrians in real time in order to make driving decisions. Given an input image, digital video, or camera images such as histology and radiology images allow for real-time many-object detection and tracking. In video analysis, facial recognition and the medical use of image classification and image segmentation are effective applications. [34]

4.2. Recurrent Neural Networks

Instead of encoding the inputs into a fixed-length vector representation, recurrent neural networks (RNNs) are an alternative approach for sequence data. RNNs are networks with loops, allowing the representation, or memory, of previous inputs. Thus, RNNs are suitable for tasks that are based on input context, like language and time-series data [35]. This makes RNNs well suited for tasks such as captioning and video, which consist of multiple frames of related visual data. In the case where the motion of the video is also an important factor, a sequence of video frames is used to reflect that motion. [36]

One of the challenges that RNNs face is that the network is short-lived, i.e., it can only recognize inputs from a few steps back. In addition, RNNs suffer from the gradient vanishing problem, making it difficult to propagate long-range dependencies. To address these challenges, RNNs have been extended with a few architectural variations to increase recognition capabilities [37]. One popular extension is the long short-term memory (LSTM) network, which introduces memory cells and a constant error flow to enable networks to determine whether input components have a short-term or long-term dependency. Another variation is gated recurrent unit (GRU), where the relationship between states occurs in its entirety within the cells. In the realm of real-world RNN applications, the combination of CNNs to extract spatial representation and RNNs to model the temporal behavior has shown success in image recognition tasks, especially in the case of sequence input.[38]

In the case of real-time action recognition, video frames are represented in a single sequence instead of spatially arranged feature maps, requiring RNN technology to understand the content contained within the video sequence. Video anomaly detection also benefits from RNN technology as it models temporal patterns for normal activities, so that situations with more significant differences than normal activities can be classified. Another application of RNN technology is hand gesture recognition. RNN technology here can recognize temporal context, making the time-local information of the movements important to approach this task.[39]

5. Applications of Image Recognition

Healthcare In the healthcare sector, image recognition technologies are combined with AI to provide medical imaging using individual pictures or X-rays to diagnose diseases. The major application areas include diagnostics and medical imaging. Picture archiving communication systems (PACSs) have the capability of providing users access to informative and high-quality pictures with a referral card [40]. The greater impact of picture archiving communication systems can be seen when information systems, such as a hospital information system, a radiology information system, a clinical information system, insurance information systems, and an integrated hospital information system were linked to it. Other applications of picture archiving communication systems include picture management, filmless radiology, telemedicine, teaching files, MSK radiographs and direct printing modalities.[41]. **2** Security Image recognition is crucial for enhancing security in various sectors, including facial recognition, surveillance systems, and alarm technologies. **3** Emerging Application Areas Image recognition will soon be integrated into food chains for all food items, focusing on accuracy standards and customer-based algorithms. The rise of critical care for patients has led to innovations aimed at reducing agricultural losses through science and technology [42]. Picture archiving communication systems facilitate swift sharing of reports. In agriculture, image recognition can monitor soil, crops, and machinery, though studying fruit recognition remains nascent and is dependent on IoT research. The effectiveness of image recognition in agriculture is under scrutiny, evaluating its reliability, benefits, and challenges [43]. Ethical considerations are vital in automating processes; developers of machine learning and AI must prioritize the ethical implications of their algorithms. While the digital society offers vast advantages, it also creates opportunities for criminal activity. Privacy issues and the high cost of personal facial recognition systems impede progress. A significant concern is the equal error rate in facial verification systems, where the likelihood of true non-matches equals false matches, potentially compromising security. Building trust between public and private sectors can mitigate political resistance to facial recognition technology. These concerns necessitate robust ethical frameworks. Facial and image recognition is now ubiquitous, identifying individuals in banks, airports, schools, and online platforms. Advertisers utilize this technology to recognize and retain customers. Visual detection aids in mitigating fraud and enhancing accuracy in ID authentication. Although these systems are powerful in various applications, their reliability remains an open question. [44]

5.1. Medical Imaging

The changes brought about by image recognition have been quite drastic in medical imaging and healthcare. A wide range of studies have shown that the automated analysis of radiographical images based on image processing algorithms, as well as modern machine learning systems and artificial intelligence, is expanding the potential of image-based diagnostic techniques, for example, in the form of visible X-ray or MRI images. The application of image processing algorithms for potential disease detection has shown promising results[45]. The application of convolutional neural networks in the case of mammograms organized into classes agreed 76% of the time with clinical annotations. These studies with new clinical applications suggest opportunities for deeper insights and are beginning to point to new interdisciplinary best practices that researchers can use for experimental tasks in which the ground truth labels are less well known. [46]

The interplay of medical image recognition, the principles of machine learning, and the connectivity to workflows has therefore been and continues to be a possible revolution for certain healthcare delivery processes. This could include improving Alzheimer's diagnosis by fusing morphometry and rest scans to identify susceptibility to developing the disease. Furthermore, supervised machine learning, a subfield of AI, has come to dominate radiomics and medical image analysis more broadly [47]. Some of the computational systems at play incorporate information extracted from the quantifiable visual features of medical images in the context of routine healthcare. Finally, continuous learning strategies for machine learning and AI models should be considered for medical image classification beyond diagnostic applications. These emerging field practices in medical research are all built upon the application of the principles of image recognition algorithms.[48]

5.2. Security and Surveillance

Introduction The safety and security of public spaces are of expanding concern in the modern era. The potential benefits of image recognition in this field could be observed in powerful surveillance and security systems. Here, we will show how computer vision and image recognition are making our living sphere safer every day [49]. An area that is most frequently associated with surveillance and security is border control and tracking threats crossing the borders. Facial recognition technology is widely used in airports and other transit locations. The possibility of this system to positively identify potential threats has saved many potential victims from danger. It can be said that such surveillance has saved many lives. [50]

Advances in the field Most of the image recognition in video surveillance is reducing the amount of crime in cities. Research on the city of London showed a reduction of crime by 20% after the installation of 10,500 cameras in the city centre. Stuttgart police in Germany also reported a 20% fall in crime after the installation of the surveillance system [51]. Another study on five French cities with population densities similar to London, which may have had 2,000 cameras instead of 10,500, discovered no significant effect in reducing crime. A case in point is the integration of deep learning systems into surveillance cameras. In the police districts of Los Angeles, a gunshot location system was operational, in which sensors embedded in public cameras send the location of the shot to the department's Real Time Analysis and Critical Response Division, which uses artificial intelligence and image recognition algorithms running in the cloud to analyses the video. [52]

6.Performance Evaluation and Comparison

Performance evaluation and comparison metrics for performance evaluation and comparison can be applied to enhance the performance of an algorithm by using the results obtained from the available image dataset. The important metrics used to assess the effectiveness of an image recognition algorithm include accuracy, precision, recall, F1-score, and confusion matrix. The accuracy of a classifier measures the frequency of instances that are correctly classified in the testing dataset. Precision is calculated to determine the percentage of predicted positive labels that are accurate. The recall measures the correct instances classified as positive by the classifier. The F1-score is a measure used to indicate the effectiveness of retrieval. The confusion matrix shows a summary of correct and incorrect classification results on a classification problem.

Performance evaluation approaches for image domain algorithms vary significantly and rely on available datasets, which act as competitions and benchmarks for comparative studies. Selecting the right dataset is crucial to ensure credible results, influencing the accuracy of the AVI system's performance. Research findings assist in identifying the most promising algorithms for specific applications. In image recognition systems, gathering a diverse training dataset is vital for classifier selection. Diversity exists in natural and synthetic datasets used for training and testing. The targets and limitations of image recognition testing depend on these datasets. To enhance performance evaluation generalization, addressing overfitting issues during testing is essential. Overfitting occurs when performance expectations exceed actual outcomes. Various strategies for performance evaluation, such as tools, techniques, classifier combinations, measures, confidence factors, and data quality thresholds, have been explored. This section elaborates on performance validation and comparison methods. It offers a comprehensive understanding of survey results related to performance evaluation and reiterates the importance of computational costs and algorithm complexity. This enables users to make informed algorithm choices, including comparisons between traditional methods, deep learning, and shallow approaches, with detailed results outlined in later sections.

7. A Comprehensive Analysis of Algorithms for Deep Learning and Machine Learning

Ref	Algorithm	Classification	Benefits	Limitations	Optimal Use Cases	Measures of Performance	Example Datasets
[19]	Naive Bayes	Traditional	Quick, comprehensible, and efficient for categorization jobs involving	Is unable to handle intricate data patterns and assumes feature independence.	Simple spam detection and image classification.	Accuracy ~80% for simple data.	Returns, 20 Newsgroups.

			text.				
[21]	Decision Trees	Traditional	Interpretable recognizes the significance of features and manages a variety of data kinds .	less stable and more prone to overfitting when data marginally changes.	exploratory study for a straightforward categorization issue.	Accuracy varies from 75 to 85% based on depth.	Boston housing and the Titanic.
[23]	Support Vector Machine (SVM)	Traditional	properly and efficiently manages high-dimensional data for binary categorization.	Large datasets are inefficient and susceptible to noise.	fraud detection with medical imaging.	F1-score>0.9 indicates high accuracy in binary tasks.	MNIST, UCIML Repository.
[24]	Random Forest	Traditional	resilient against noise and uses ensemble methods to reduce overfitting.	less interpretable than a single tree and computationally costly.	Multi-class categorization and feature selection.	Accuracy >90%, robust F1-score.	MNIST, Kaggle Datasets.
[26]	k-nearest neighbors (KNN)	Traditional	Easy to use, adaptable, and efficient for small datasets.	sensitive to unimportant factors and computationally demanding for huge data sets.	face recognition and optical character recognition (OCR).	High accuracy (>85%) for small datasets.	Iris, CIAR-10.
[31]	Convolution Neural Networks (CNN).	Deep Learning	Automated feature extraction performs exceptionally well in object detection and image categorizati	demands a lot of data and processing power.	security systems, medical imaging, and driverless cars.	On big datasets, accuracy is greater than 95%.	Image Net, COCO

			on. .				
[33]	Adversarial Generative Networks	Deep Learning	produces lifelike visuals that can be used to enhance data.	unstable (mode collapse) and necessitates rigorous training.	Generating synthetic data and transferring styles .	FID<50, quality of perception changes.	Celeb A, LSUN.
[35]	Recurrent Neural Networks (RNNs)	Deep Learning	efficiently captures temporal correlations through the processing of sequential data.	has limited memory and fading gradients. .	gesture recognition and video analysis.	BLEU score for sequence problems is greater than 30, and performance varies.	Kinetics, Charades.
[38]	Transformers	Deep learning	performs well on problems involving several modes and sequences.	significant training requirements and a high computational expense.	Multimodal analysis of the image comments	>35 BLEU, >90 % multimodal accuracy.	MS COCO, Flickr 30K
[45]	YOLO (You Only Look Once	Deep learning	fast and accurate real-time object detection.	finds it difficult to identify little objects in cluttered environments.	Autonomous cars and security.	Map>50%, real-time FPS>30.	COCO, Pascal VOC
[48]	RESET	Deep learning	allows for deep network training and resolves vanishing gradient problems.	It is computationally intensive and needs a lot of tagged data.	Transfer learning and feature extraction.	Top-5 accuracy>95 % on ImageNet.	ImageNet, CIFAR-10.

8. SUMMARY AND OUTLOOK

Summary. Image recognition is conducted through two groups of algorithms: traditional methods and deep learning approaches. Traditional methods call for predefined features, such as the histogram of oriented gradients for edge detection. However, features selected in this fashion are artificial and may degrade the recognition performance. Deep learning was adopted with the intent of learning features from an image based on the massive labeled data. This amounts to developing a type of artificial neural network. It has achieved several milestones, including the use of a voting system of a set of learned machines to reduce the error rate and the improvement in the efficiency of deep learners by introducing technologies such as parallel computing. The development of deep learning, as distinguished from classic learning, is based on a large amount of image data and overcomes the problem of feature abstraction. In recent years, the related work on image recognition has continued to evolve in the pursuit of even greater performance; however, determining the problems to be solved and implementing a resolution remain worthwhile pursuits. In recent years, especially in biology, electronics, medicine, and other related fields, image recognition has become an essential facet and attracted widespread attention. In addition to biological recognition, there are numerous practical applications such as the identification of iris and fingerprint, face alignment systems in cameras, house surveillance for home security, and monitoring systems for the environment, which support the conclusion that image recognition technology can greatly improve efficiency. Potential future studies can strive to acknowledge further deep learners that may exist for image recognition, enhance the accuracy and efficiency of large-scale image recognition, address face identification for underexposed-to-daylight variations, and contribute further discussion regarding the moral implications of image recognition technologies. Furthermore, establishing a responsible strategy for the deployment of this image recognition technology is also deserving of consideration. Moreover, future studies must adapt and train the image recognition algorithms to confront emerging challenges. It is hoped that the present essay facilitates future scientific communications focusing on these branches, as mentioned above.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

References:

[1] X. Wu, D. Sahoo, and S. C. H. Hoi, "Recent advances in deep learning for object detection," Neurocomputing, 2020. [PDF]

[2] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: a new paradigm to machine learning," Archives of Computational, 2020. academia.edu

[3] L. F. Li, X. Wang, W. J. Hu, N. N. Xiong et al., "Deep learning in skin disease image recognition: A review," Ieee Access, 2020. ieee.org

[4] M. Soori, B. Arezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics, a review," Cognitive Robotics, 2023. sciencedirect.com

[5] Z. Zhang, F. Wen, Z. Sun, X. Guo, and T. He, "Artificial intelligence□enabled sensing technologies in the 5G/internet of things era: from virtual reality/augmented reality to the digital twin," Advanced Intelligent, 2022. wiley.com

[6] A. M. Akanksh, M. S. Nayak, and S. Nishith, "Automated invoice data extraction using image processing," Journal of Artificial, 2023. [HTML]

[7] H. Ajra, M. K. Nahar, and L. Sarkar, "Disease detection of plant leaf using image processing and CNN with preventive measures," in Emerging Technology in ..., 2020. [HTML]

[8] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," SN computer science, 2021. springer.com

[9] A. Venon, Y. Dupuis, and P. Vasseur, "Millimeter wave FMCW radars for perception, recognition and localization in automotive applications: A survey," in IEEE Transactions on ..., 2022. [HTML]

[10] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition–A review," Information Processing in ..., 2021. sciencedirect.com

[11] M. Salvi, U. R. Acharya, and F. Molinari, "The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," Computers in Biology and Medicine, 2021. sciencedirect.com

[12] C. L. Chowdhary and D. P. Acharjya, "Segmentation and feature extraction in medical imaging: a systematic review," Procedia Computer Science, 2020. sciencedirect.com

[13] V. K. Vishnoi, K. Kumar, and B. Kumar, "Plant disease detection using computational intelligence and image processing," Journal of Plant Diseases and Protection, 2021. [HTML]

[14] F. Piccialli, V. Di Somma, F. Giampaolo, and S. Cuomo, "A survey on deep learning in medicine: Why, how and when?," Information, Elsevier, 2021. [HTML]

[15] FZ Canal, TR Müller, JC Matias, and GG Scotton, "A survey on facial emotion recognition techniques: A state-of-the-art literature review," Information, Elsevier, 2022. [HTML]

[16] Y. Liu, H. Pu, and D. W. Sun, "Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices," Trends in Food Science & Technology, 2021. [HTML]

[17] S. Mascarenhas and M. Agarwal, "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification," in 2021 International Conference on ..., 2021. [HTML]

[18] R. Zebari, A. Abdulazeez, D. Zeebaree, and D. Zebari, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," Journal of Applied, 2020. jastt.org

[19] N. O'Mahony, S. Campbell, A. Carvalho, "Deep learning vs. traditional computer vision," Advances in Computer, Springer, 2020. [PDF]

[20] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," Applied Sciences, 2022. mdpi.com

[21] L. Chen, S. Li, Q. Bai, J. Yang et al., "Review of image classification algorithms based on convolutional neural networks," Remote Sensing, 2021. mdpi.com

[22] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," International Journal of …, 2022. [HTML]

[23] D. A. Pisner and D. M. Schnyer, "Support vector machine," Machine learning, 2020. [HTML]

[24] M. Sheykhmousa and M. Mahdianpari, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," IEEE Journal of, 2020. ieee.org

[25] Z. Jun, "The development and application of support vector machine," Journal of Physics: Conference Series, 2021. iop.org

[26] Z. Pan, Y. Wang, and Y. Pan, "A new locally adaptive k-nearest neighbor algorithm based on discrimination class," Knowledge-Based Systems, 2020. [HTML]

[27] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers-a tutorial," ACM computing surveys (CSUR), 2021. acm.org

[28] M. Capra, B. Bussolino, A. Marchisio, G. Masera, "Hardware and software optimizations for accelerating deep neural networks: Survey of current trends, challenges, and the road ahead," IEEE, 2020. ieee.org

[29] G. Menghani, "Efficient deep learning: A survey on making deep learning models smaller, faster, and better," ACM Computing Surveys, 2023. [PDF]

[30] K. Berggren, Q. Xia, K. K. Likharev, D. B. Strukov, "Roadmap on emerging hardware and technology for machine learning," 2020. iop.org

[31] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artificial intelligence review, 2020. [PDF]

[32] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," ISPRS Journal of Photogrammetry and Remote Sensing, 2021. academia.edu

[33] H. Garg, B. Sharma, S. Shekhar, and R. Agarwal, "Spoofing detection system for e-health digital twin using EfficientNet Convolution Neural Network," in Tools and Applications, Springer, 2022. [HTML]

[34] S. S. Kshatri and D. Singh, "Convolutional neural network in medical image analysis: a review," Archives of Computational Methods in Engineering, 2023. [HTML]

[35] P. B. Weerakody, K. W. Wong, G. Wang, and W. Ela, "A review of irregular time series data handling with gated recurrent neural networks," Neurocomputing, 2021. [HTML]

[36] A. F. Ansari, L. Stella, C. Turkmen, and X. Zhang, "Chronos: Learning the language of time series," arXiv preprint arXiv:2024. [PDF]

[37] W. Ullah, A. Ullah, I. U. Haq, K. Muhammad, "CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks," Multimedia Tools and Applications, Springer, 2021. academia.edu

[38] W. Ullah, A. Ullah, T. Hussain, K. Muhammad, "Artificial Intelligence of Things-assisted two-stream neural network for anomaly detection in surveillance Big Video Data," Future Generation, Elsevier, 2022. [HTML]

[39] W. Ullah, A. Ullah, T. Hussain, Z. A. Khan et al., "An efficient anomaly recognition framework using an attention residual LSTM in surveillance videos," Sensors, 2021. mdpi.com

[40] I. Castiglioni, L. Rundo, M. Codari, G. Di Leo, and C. Salvatore, "AI applications to medical images: From machine learning to deep learning," Physica Medica, Elsevier, 2021. physicamedica.com

[41] A. S. Panayides, A. Amini, N. D. Filipovic, "AI in medical imaging informatics: current challenges and future directions," … and health …, 2020. ieee.org

[42] S.K. Zhou, H. Greenspan, C. Davatzikos, et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," Proceedings of the IEEE, 2021. nih.gov

[43] M. Ghiasi, T. Niknam, Z. Wang, and M. Mehrandezh, "A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future," Electric Power Systems, Elsevier, 2023. [HTML]

[44] S. Bhattacharya, P. K. R. Maddikunta, and Q. V. Pham, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey," Sustainable Cities and ..., 2021. nih.gov

[45] V. K. Singh, H. A. Rashwan, S. Romani, F. Akram, "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," in Applications, Elsevier, 2020. [PDF]

[46] M. A. Hernández-Vázquez, Y. M. Hernández-Rodríguez, "Hybrid Feature Mammogram Analysis: Detecting and Localizing Microcalcifications Combining Gabor, Prewitt, GLCM Features, and Top Hat Filtering," Diagnostics, 2024. mdpi.com

[47] X. Chen, X. Wang, K. Zhang, K. M. Fung, and T. C. Thai, "Recent advances and clinical applications of deep learning in medical image analysis," in Medical Image Analysis, Elsevier, 2022. sciencedirect.com

[48] J. Wang, H. Zhu, S. H. Wang, and Y. D. Zhang, "A review of deep learning on medical image analysis," Mobile Networks and Applications, 2021. [HTML]

[49] M. Farhadmanesh, A. Rashidi, and A. K. Subedi, "A computer vision-based standalone system for automated operational data collection at non-towered airports," in IEEE, 2024. ieee.org

[50] M. Smith and S. Miller, "The ethical application of biometric facial recognition technology," Ai & Society, 2022. springer.com

[51] C. Hein and D. Schubert, "Resilience and path dependence: A comparative study of the port cities of London, Hamburg, and Philadelphia," Journal of Urban history, 2021. sagepub.com

[52] A. E. Nivette, R. Zahnow, R. Aguilar, A. Ahven, "A global analysis of the impact of COVID-19 stay-at-home restrictions on crime," Nature Human Behavior, 2021. nature.com