Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 8, No. 4, November, Pp. 759-772

A Systematic Review of Artificial Intelligence in Assistive Technology for People with Visual Impairment

Liliek Triyono[1,2*], Rahmat Gernowo[1], Prayitno[2], Saifur Rohman Cholil[1,3]
Doctoral Information System, Diponegoro University, Semarang, Indonesia[1]
Department of Electrical Engineering, Politeknik Negeri Semarang, Semarang, Indonesia[2]
Departement of Information System, Universitas Semarang, Semarang , Indonesia[3]

Article Info

Keywords:
Deep Learning; Artificial Intelligence; Visual Impairment; Visual Aids; Assistive Technology

Article history:
Received: June 29, 2023
Accepted: September25, 2023
Published: November 30, 2023

Cite:
L. Triyono, R. Gernowo, Prayitno, and S. R. Cholil, "A Systematic Review of Artificial Intelligence in Assistive Technology for People with Visual Impairment", KINETIK, vol. 8, no. 4`, Nov. 2023. https://doi.org/10.22219/kinetik.v8i4.1772

*Corresponding author.
Liliek Triyono
E-mail address:
liliektriyono@students.undip.ac.id

Abstract

Recent advances in artificial intelligence (AI) have led to the development of numerous successful applications that utilize data to significantly enhance the quality of life for people with visual impairment. AI technology has the potential to further improve the lives of visually impaired individuals. However, accurately measuring the development of visual aids continues to be challenging. As an AI model is trained on larger and more diverse datasets, its performance becomes increasingly robust and applicable to a variety of scenarios. In the field of visual impairment, deep learning techniques have emerged as a solution to previous challenges associated with AI models. In this article, we provide a comprehensive and up-to-date review of recent research on the development of AI-powered visual aides tailored to the requirements of individuals with visual impairment. We adopt the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology, meticulously gathering and appraising pertinent literature culled from diverse databases. A rigorous selection process was undertaken, appraising articles against precise inclusion and exclusion criteria. Our meticulous search yielded a trove of 322 articles, and after diligent scrutiny, 12 studies were deemed suitable for inclusion in the ultimate analysis. The study's primary objective is to investigate the application of AI techniques to the creation of intelligent devices that aid visually impaired individuals in their daily lives. We identified a number of potential obstacles that researchers and developers in the field of visual impairment applications might encounter. In addition, opportunities for future research and advancements in AI-driven visual aides are discussed. This review seeks to provide valuable insights into the advancements, possibilities, and challenges in the development and implementation of AI technology for people with visual impairment. By examining the current state of the field and designating areas for future research, we expect to contribute to the ongoing progress of improving the lives of visually impaired individuals through the use of AI-powered visual aids.

1. Introduction

The eye is the primary visual organ in the human body [1,2]. To receive external light stimuli, the eyes must come into contact with external space, which increases their likelihood of coming into contact with external microorganisms. Some viruses can invade the body through the eyes, resulting in eye diseases and impaired vision [3–5]. By 2019, the World Health Organization [6] estimates that at least 220 million people will have visual impairment.

Blind individuals [7–9] are blind in one or both eyes due to a disease or incidental injury to the optic nerve or eye. Contrary to a common misconception, not all individuals with visual impairment are blind; blindness is just one form of visual impairment. Visual impairment encompasses a spectrum of conditions, and some individuals with visual impairment retain the ability to see objects to varying degrees. Visual impairment [10] is a condition characterized by impaired vision, frequently manifested in differing degrees of near or farsightedness. Color blindness [11]–[13], reduced visual acuity [14], night blindness [15]–[17], visual field defect [18]–[21], double vision [22], and black silhouettes in front of the eyes are visual impairment symptoms. Visual impairment is not a singular ophthalmic condition, but can be caused by a variety of conditions. Refractive error, cataracts, funduscopic disease, and optic neuropathy are common diseases that can result in visual impairment.

Due to their physical disability, individuals with visual impairment experience numerous inconveniences in daily life [23]. When they go outside, they cannot see impending traffic or changes in traffic signals and cannot avoid obstacles in time. To enhance their personal development through education, individuals with visual impairment are encouraged to devote considerable effort to acquiring proficiency in braille, as well as reading and writing skills. These skills play a vital role in empowering visually impaired individuals to access information, engage in learning activities, and expand their educational horizons. In addition to navigating daily life, visually impaired individuals frequently have difficulty gaining access to necessary medical care. They rely on their senses of touch and hearing as well as assistance to live a normal existence.

The topic of artificial intelligence is currently very popular all over the globe. One of the key themes for the future is the intelligence of things [24]–[27]. Big data and cloud computing are helping to advance artificial intelligence (AI) in a number of areas. Intelligent inventory management is more precise and quick than manual management in the retail and logistics industries [28], [29]. It can also analyze sales patterns and take part in forecasting upcoming trends in merchandise sales [30].

Customers can streamline the purchasing process, obtain tailored product recommendations, and enhance their experience with artificial intelligence. Planning routes from the source to the target can be aided by artificial intelligence when it comes to shipping goods [31], [32]. Emerging AI-based educational tools that track students learning processes and customize them to their unique requirements can be used to supplement conventional educational methods [33], [34]. Using its strengths in image processing and vehicle location in surveillance videos, artificial intelligence in security can assist police personnel in finding suspects in surveillance videos [35], [36]. The artificial intelligence-based system can compile swipe records [37], [38] and tracking data [39], [40] for the entire building and alert of anomalies when they happen.

Between 2019 and 2023, significant advancements in the field of artificial intelligence led to the publication of several noteworthy studies in reputable journals. This accelerated development necessitates an updated examination of the use of artificial intelligence in assisting individuals with visual impairment. This paper's primary objective is to provide a comprehensive overview of recent developments, such as the examination of vision testing criteria, the application of individual models in deep learning for diagnosing and classifying eye diseases, and the emergence of artificial intelligence wearable devices designed to assist individuals with visual impairment.

This study differs from existing reviews. General descriptions of AI are given in [25]–[27], while detailed discussions of recent challenges are presented in [33], [34], tracking data [39], [40]. Resumes of AI applications in dilated fundus examinations [41] and small pupil fundus examinations [42], [43] which focus on the vitreous [44], retina [45], choroid [46], and optic nerve [47], [48] also have been published. Despite the existing body of research on artificial intelligence (AI) in the context of visual impairment, there is a notable lack of in-depth analysis regarding the performance and impact of AI. Additionally, there is a need for a comprehensive examination of benchmarking practices in AI for visual impairment technology. To address these gaps, this review aims to provide a thorough survey encompassing various aspects, including the application of AI in diagnosing ocular diseases, the development of smart devices for assisting visually impaired individuals, benchmarking methodologies, and the promising potential applications of AI in this field.

In essence, the main aim of this scholarly article is to address the current knowledge gaps pertaining to the utilization of artificial intelligence for the purpose of aiding those who experience visual impairment. The objective of this study is to thoroughly investigate recent developments in vision testing standards, the utilization of specific deep learning models for the purpose of diagnosing and categorizing eye ailments, and the emergence of wearable artificial intelligence devices intended to enhance the quality of life for individuals with visual impairments. Through the pursuit of these aims, our aim is to provide significant contributions to the area, with the ultimate goal of improving accessibility and quality of life for those with visual impairment.

2. Research Method

This study was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [49]. Adopted by health-related organizations and journals, the PRISMA technique is a standard for reporting evidence in systematic reviews that is widely recognized [50]. PRISMA approaches offer numerous benefits, including highlighting the review's quality, enabling readers to evaluate the review's strengths and weaknesses, replicating review procedures, and structuring and formatting the review using PRISMA headings [49]. However, conducting a systematic review and publishing it thoroughly can be time-consuming. In addition, it can quickly become obsolete; therefore, it must be routinely updated to include all newly published primary sources since the project's inception.

2.1 Formulate Research Questions

The research query is divided into the following sub-questions.

- RQ1: What are the most advanced AI techniques in the visual impairment?
- RQ2: What are the most recent datasets used in visual impairment?
- RQ3: What are the research gaps and prospective future research directions related to visual impairment application research?

The objective of the first research question (RQ1) is to provide an exhaustive and systematic overview of all AI-related articles. In addition, RQ1 seeks to provide evidence that the field of visual impairment can benefit from the incorporation of AI. In addition, the second research question's (RQ2) motivation is to address AI-settings challenges such as the datasets and the development of intelligent devices for the visually impaired. Lastly, the third research question (RQ3) presents prospects for AI researchers, with a focus on application difficulties for the visually impaired.

2.2 Data Eligibility and Analysis of the Literature

Figure 1 depicts the PRISMA flowchart [48] used in the article selection process, which outlines the search, inclusion, and exclusion of papers. The PRISMA flowchart contains three steps: identification, screening, and inclusion. First, in the identification phase, we conducted a comprehensive literature review between 1 January 2019 and 30 March 2023 utilising Springer, IEEE Xplore, the Association for Computing Machinery Digital Library (ACM DL), Science Direct, and Scopus. We begin in 2019 because we are interested in further implementation in visual impairment field at last 5 years before. In general, the following search terms were utilised: "Artificial Intellegence", "Visual Impairment", and "Assistive Technology". Due to the fact that each publication database has its own set of search query filters, the specific query terms are detailed in Table A1 of Appendix A. Initial search results from digital libraries revealed 322 articles that met the search criteria. Then, 2 duplicate articles were removed, leaving 320 articles in the identification phase.

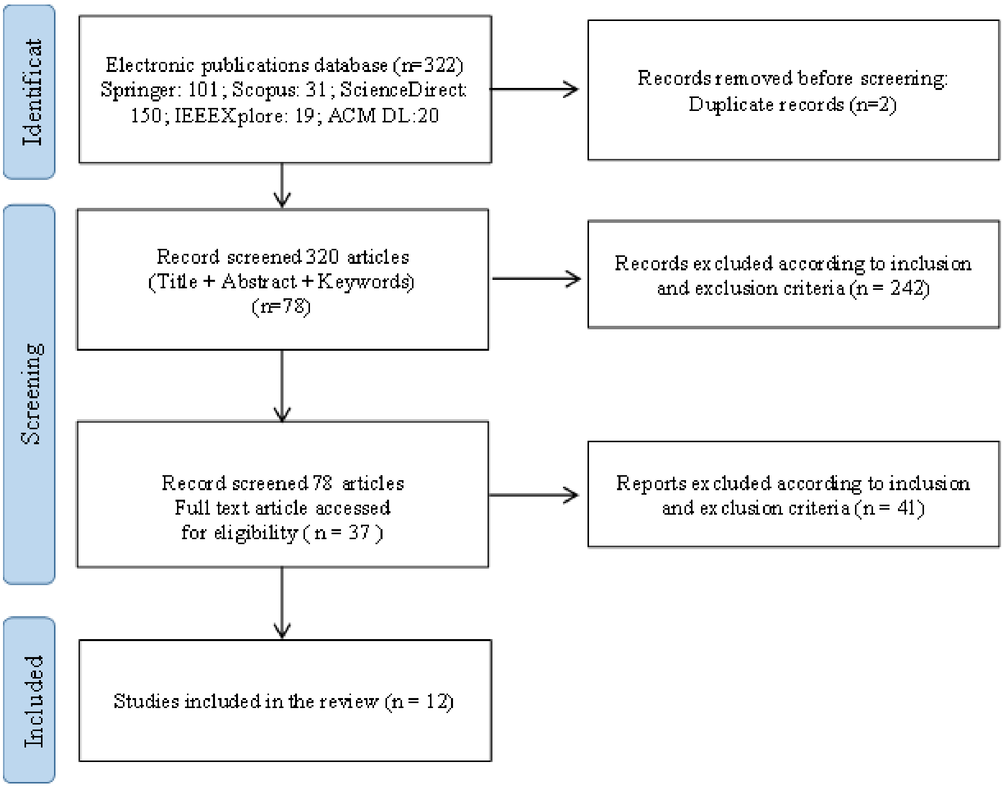

Figure 1. Study Selection using PRISMA Flow Diagram Method Consisting of Identification Step, Screening Step, and Included Step

While systematic reviews offer numerous benefits, they are susceptible to biases that obscure the objective results of the study and should be evaluated with caution [51]. To ensure the research selection process was unbiased and free from ambiguity, several methods were implemented. These methods included (i) conducting a dual review, (ii) establishing explicit and transparent inclusion and exclusion criteria, and (iii) documenting the flow of the selection process using the PRISMA flow diagram. Following this, both the abstracts and full texts of relevant articles were meticulously examined, and only those that met the predetermined inclusion and exclusion criteria were chosen. Additionally, researchers cross-verified the selected papers and resolved any discrepancies or conflicts through consensus. There was no disagreement regarding the documents contained in this review.

This study provides an overview of AI in the field of visual impairment and should delve deeper into setting secure mechanisms for AI medical data. Therefore, the screening step defines inclusion and exclusion criteria. We included publications that (i) use AI to develop models on image-based datasets, (ii) are published in prominent peer-reviewed journals, and (iii) are published in English. Exclusion criteria were used to exclude irrelevant published studies based on the following criteria: (i) articles not related to AI, (ii) non-visually impaired or experimentally image-based datasets, (iii) languages other than English, (iv) reviews, (v) proceedings or conference papers, and (vi) arXiv preprints/section.

A number of considerations oppose the inclusion of conference papers in this study. First, conference proceedings typically cover a wide variety of topics and larger publications, so finding the right conference, accessing the abstracts, and searching thousands of abstracts can be time-consuming and can be resource intensive. Second, conference reports may not contain enough information for systematic reviewers to assess methods, risk of bias, and results of studies submitted to conferences because of their brevity. Finally, the reliability of results, especially in the health field, has also been questioned. One reason is that the results are often preliminary or based on limited research to meet conference deadlines. Therefore, conference papers are not included in the selection criteria.

After implementing the criteria for inclusion and exclusion to the titles, abstracts, and keywords of each study, 78 articles were identified during the screening process. Next, the 37 articles in the report that were evaluated in the eligibility step were then excluded from the full article text by the exclusion criteria, resulting in 12 articles. Finally, in the included steps, 12 articles using AI in visual impairment applications were selected for further analysis, the results of which are described in this study.

To offer a numerical summary of the available literature search, the following data was extracted from each article: (i) Author, Title, Year, and Keywords; (ii) proposed methods including training artificial intelligence algorithms and deep learning/machine learning models; (iii) data characteristics including distribution methods and challenges; and (iv) Test Results and Discussion.

3. Results and Discussion

All researchers labored diligently on this comprehensive literature review. Following the completion of the preceding stages, the papers were distributed among the researchers. Then, each researcher thoroughly read their assigned papers and categorized them into themes based on the subject they addressed. Researchers held multiple meetings to discuss and define these topics. The researchers determined the themes based on the issues discussed in the reviewed papers. In the conclusion, the researchers identified seven themes. The references chosen for each topic are listed in Table A1. Each researcher was then designated a theme to summarize their investigation and report the results. In this section, the results are discussed.

Numerical description. The following observations are based on a quantitative examination of the 12 included studies between 2019 and June 2023. First, the quantity of AI studies published in medical journals by publication year. Since the year 2000, the number of articles published on AI has increased steadily. The number of published papers in 2022 should increase linearly throughout the year.

Deep learning algoritms. In addition, we wish to describe the deep learning models used in the experiments and assess their proposed visual impairment algorithms.

The strengths and limitations of artificial intelligent algorithms for visual impairment are summarized in Table 1.

Table 1. Summary of Artificial Intelligent Algorithms Performing on Visual Impairment, along with Strengths and Weaknesses.

AI Algorithms	Strength	Weakness	AI Study
AE	By utilizing unsupervised learning, AE is primarily intended for dimensional feature reduction and denoising of medical datasets. AE seeks to recreate a concise and effective representation of features.	An auto encoder may omit vital information from the characteristics of an object detection dataset.	[52]
CNN	Excellent performance on object detection image classification tasks such as scene environtment prediction using X-ray images.	If the client in the FL environment lacks potent computational resources, the training procedure of a CNN with multiple layers will be time-consuming.	[53]
GAN	Generate a synthetic sample of medical data for	Training GAN is difficult because the training process	[54]

	experiments with limited data sets.	is unstable, there is no standard metric evaluation, and numerous trial-and-error experiments are required for effective results.	
LSTM	Excellent performance on time series or sequential medical datasets, such as human activity recognition detection.	Training LSTM is difficult due to the difficulties of vanishing and explosive gradients.	[55][56]
MLP	Excellent generalization performance on tabular common object datasets, such as indoor goods.	MLP is limited to rudimentary problem solving. In addition, it is sensitive to feature scaling and requires the configuration of numerous hyperparameters, including the number of concealed neurons and layers.	[57][58]
SVM	SVM can model nonlinear decision boundaries, and numerous kernels are available. In addition, it is resistant to overfitting, especially in high-dimensional space.	SVM is memory-intensive, harder to modify due to the critical selection of the appropriate kernel, and does not extend well to larger datasets.	[59][60]
U-Net	Obtain precise results when segmenting medical image datasets, such as when segmenting brain tumors disease using brain magnetic resonance medical images.	Development of the U-Net architecture is time-consuming because the network must be operated independently for each patch, and because overlapping patches create redundancy. In addition, there exists a compromise between the precision of localization and the use of context.	[61]
ViT	ViTs are highly scalable. They can handle both small and large input images without requiring major architectural changes	Training large Vision Transformer models can be computationally expensive and time-consuming, making them less accessible to researchers with limited resources.	[62][63]

AE: autoencoder; CNN: convolutional neural network; GA.: generative adversarial network; LSTM: long short-term memory; MLP: multilayer perceptron; SVM: support vector machine; ViT: Vision Transformer.

.

RQ1: What are the most advanced AI techniques in the visual impairment?

3.1 Recent Methods of AI

With enormous technological advancements in computer vision and deep learning, artificial intelligence is increasingly being used in medical imaging [64]–[67]. As a novel treatment modality, this is gaining more interest. The incorporation of AI techniques into the treatment of visual impairment has improved the accuracy and speed of clinical screening, diagnosis, and prognosis. In diagnosing the eyes, fundus images are readily available and contain abundant biological data, making them suitable for CNN, GAN, and transfer learning.

All three methods are effective at image analysis processing. CNNs can directly convolve with image pixels to extract image features. GANs can generate new images based on the characteristics of actual data, in addition to the classification and feature extraction capabilities of conventional neural networks. It is not necessary to presume additional mathematical hyperparameters in practice. Transfer learning reduces the requirement that training and test data be independently and identically distributed. Moreover, it eliminates the need for manual annotation of the dataset during the training procedure. Poor generalization performance, time-consuming and onerous procedures, and limited datasets can be effectively addressed through transfer learning.

3.1.1 CNN

CNN, an essential subfield of deep learning [68], is frequently applied to image and video processing. CNNs differ from conventional neural networks in that their feature-learning stage is comprised of one or more convolutional and pooling layers. Each neuron in the convolutional layer is not connected to every neuron in the subsequent layer, thereby reducing the computational burden of training a conventional fully connected neural network. Figure 2.

The extraction of features is the procedure of convolution. One kernel of convolution corresponds to one feature. Each output feature map from the preceding layer is convolved with K-convolution kernels to generate a feature map for the subsequent layer. The details can be seen in Equation 1.

$$h_j^{(n)} = \sum_{k=1}^{K} h_k^{n-1} \otimes W_{kj}^{(n)} + b_{kj}^{(n)} \tag{1}$$

Where $\otimes$ is a 2-dimensional convolution. The output of the *j*-th feature map in the *n*-th concealed layer is represented by $h_j^{(n)}$. h_k^{n-1} is the *k*-th channel in the *(n-1)*-th concealed layer. $w_{kj}^{(n)}$ denotes the weights of the *k*-th channel within the *j*-th filter of the *n*-th layer. $b_{kj}^{(n)}$ is the pertinent bias term [69].

The subsequent layer generates as many convolution kernels as feature maps. The greater the value in the feature map acquired after convolution, the greater the resemblance between the image in that region and the kernel's features.

N_{para} is presented in Equation 2 describing the number of parameters in a convolutional layer. [70]:

$$N_{para} = C_0 \times (k_w \times k_h \times C_i + 1) \tag{2}$$

The Equation 3 is the computation required in a convolutional layer [71], presented as follows [71]:

$$F = [(C_i \times k_w \times k_h) + (C_i \times k_w \times k_h - 1) + 1] \times C_o \times W \times H \tag{3}$$

where F represents the floating-point operation and C_i is the number of input channels. C_0 stands for the number of output channels, k_w for the width of the convolution kernel, k_h for the height of the convolution kernel, *W* for the length of the feature map, and *H* for its width. +1 represents bias. $C_i \times k_w \times k_h$ represents the number of multiplication operations in one convolution operation, $C_i \times k_w \times k_h - 1$ represents the number of subtraction operations in one convolution operation. $C_o \times W \times H$ indicates the total number of elements in the feature map.

To address the issue where the output image is marginally smaller than the input image, researchers may patch the image prior to the convolution operation. The required number of pixels are then added to the image's outermost layer. This mitigates the disadvantage that information from image corners or margins plays a lesser role. The structure of a convolutional neural network is depicted in Figure 2.

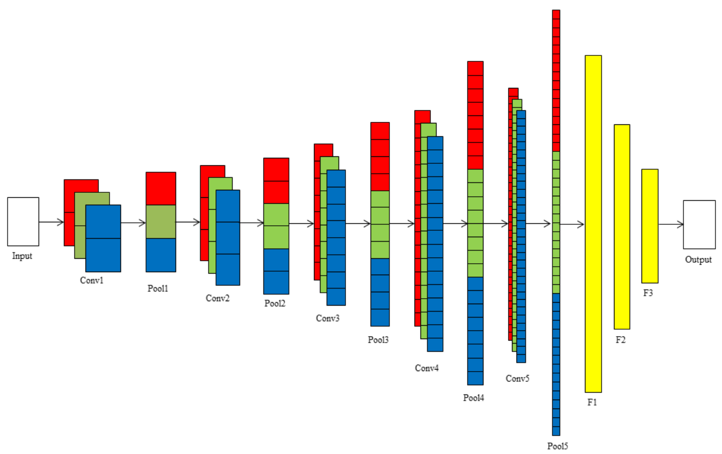

Figure. 2. The Structure of a CNN

3.1.2 GAN

The primary components of generative adversarial networks (GANs) [72], [73] are the generator and discriminator. The generator produces fictitious images based on arbitrarily generated noise. The purpose of the generator is to produce images that closely resemble those in the actual dataset. These fabricated data are combined with images from the actual dataset and then fed to the discriminator. The discriminator determines whether the image being classified is from the actual dataset or was generated by the generator based on the extracted features. It defines losses regarding the maximization of the discriminator and the minimization of the generator. The Equation 4 is:

$$\min_{G} \max_{D} V(D,G) = E_{x\ P_{data(x)}}[\log D(x)] + E_{z\ P_{z(z)}}\{\log[1 - D(G(z))]\}, \tag{4}$$

where x represents the true data, $p_{\text{data}}(x)$ is the probability distribution of *x, z* represents the noisy data, $p_z(z)$ is the probability distribution of *z*, and the generated dummy sample data *G(z)*. *D(x)* represents the probability that the discriminator will determine whether the picture is true, and *D(G(z))* is the probability that the discriminator will determine whether the picture generated by *G* is true. *E* represents the mathematical expectation [74].

During the game between the two adversarial models, both models continuously adjust their parameters to improve their ability to create and identify fake images. Figure 3 shows the workflow of the GAN model.

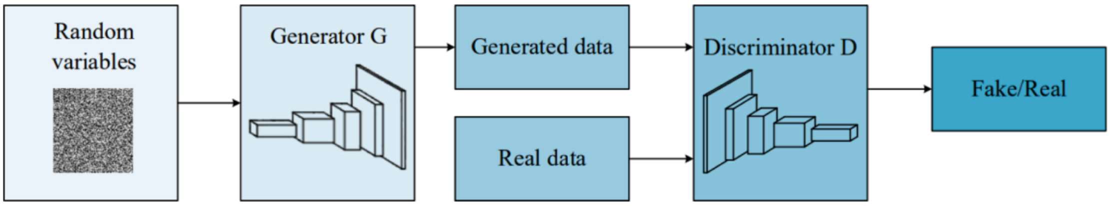

Figure. 3 The GAN Model Workflow

3.1.3 Transfer Learning

The objective of transfer learning [75], [76] is to bring the computer closer to how humans actually think than machine learning. When humans encounter a problem for the first time, they associate it with a previously encountered and resolved problem. People subconsciously compare two similar problems and attempt to derive a solution for the latter from the former. Transfer learning simulates this process by (i) applying what has been learned from previous tasks to a new problem and (ii) acquiring a superior solution by learning in a new domain.

To learn features and transfer models during training, an increasing number of researchers are integrating emerging neural networks with transfer learning methods. Transfer learning has numerous applications, including computer vision, text classification, behavioral recognition, natural language processing, indoor localization, and video surveillance, among others.

Feature extraction and fine-tuning are two transfer learning techniques typically employed. People use the representations learned by the prior network to extract meaningful features from new data and add a new classifier during feature extraction. This method requires less time because the basal convolutional network contains normally valuable features for image classification.

Fine-tuning is the process of adjusting the weights of the network's primary layers and reconfiguring the later layers as required for the task, thereby revising the weights of the later or particular layers. Lower-level features are extracted by the fundamental convolutional layers. However, as the depth of the strata increases, more distinctive characteristics are derived from subsequent layers. Lower-level characteristics are more general, whereas higher-level characteristics are more specific. The features learned in the final few layers of the model vary significantly between datasets. During training, slower learning rates are frequently employed to prevent significant changes in pre-trained parameters and to prevent over-fitting.

3.1.4 Combined Multi-technology Approaches

CNN, GAN, and transfer learning each have their own domains of application. CNNs perform significantly better in the domain of supervised learning than in the domain of unsupervised learning. Therefore, some researchers wish to combine CNN and GAN to create the deep convolutional generative adversarial network (DCGAN) [77], which performs well in the domain of unsupervised learning. CNN's incorporation into GAN bridges the divide between supervised and unsupervised learning. In contrast to GAN, it employs full convolution rather than full connection. The generator makes use of LeakyReLU. The discriminator employs the ReLU language. The bulk normalization layers are utilized by the generator and discriminator to ensure as stable a training environment as feasible. DCGAN employs the convolution step stride to control whether or not downsampling is conducted. Combining transfer learning with CNN network models is more prevalent.

The concept of integrating GAN and transfer learning entails selecting available features for transfer between domains. The method [78] permits training on a large scale using a large amount of annotated data in the source domain and a large amount of unannotated data in the target domain. When the data is mapped to a particular feature space, the labeled predictor is able to differentiate the data class from the source domain.

In contrast, the domain discriminator is unable to determine the origin of the data. For labeled data from the source domain, the network minimizes the loss of the labeled predictor in a continuous manner. For all data from the source and target domains, the network minimizes the domain discriminator loss continuously.

RQ2: What are the most recent datasets used in visual impairment?

3.2 Dataset

This Tabel 2 section provides an overview of the most widely used datasets for object detection tasks.

Table 2. Comparison of Various Object Detection Datasets

Dataset	Classes	Train			Validation			Test
		Images	Objects	Objects/Image	Images	Objects	Objects/Image	
OpenImage	600	1,743,042	14,610,229	8.38	41,620	204,621	4.92	125,436
ILSVRC	200	456,567	478,807	1.05	20,121	55,501	2.76	40,152
MS-COCO	80	118,287	860,001	7.27	5,000	36,781	7.35	40,670
PASCAL VOC 12	20	5,717	13,609	2.38	5,823	13,841	2.37	10,991

3.2.1 Pascal VOC 07/12

The Pascal Visual Object Classes (VOC) competition was a multiyear endeavor to advance the field of visual perception. It began in 2005 with classification and detection tasks on four object classes [83]; however, only two versio\ns of these challenges are typically used as a benchmark. The VOC07 challenge included 5,000 training images and more than 12,000 labeled objects [84] whereas the VOC12 challenge included 11,000 training images and more than 27,000 labeled objects [79]. The number of object classes has been increased to twenty, and tasks such as segmentation and action detection have been added. To assess the efficacy of the models, Pascal VOC introduced the mean Average Precision (mAP) at 0.5 IoU (Intersection over Union).

3.2.2 ILSVRC

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [80] was an annual competition from 2010 to 2017 that established a standard for evaluating algorithm performance. The size of the dataset was increased to more than one million images comprising one thousand object classification classes. These 200 classes were hand-selected for the object detection assignment, which consists of over 500,000 images. Multiple sources, such as ImageNet [79] and Flikr, were utilized to compile the detection dataset. ILSVRC also modified the evaluation metric by lowering the IoU threshold to facilitate the detection of tiny objects.

3.2.3 MS-COCO

Microsoft Common Objects in Context (MS-COCO) [81] is one of the most complex available datasets. It contains 91 common objects found in their natural environment that a 4-year-old can readily identify. It was introduced in 2015, and since then, its popularity has only grown. It has an average of 3.5 categories per image and more than two million instances. Additionally, it contains 7.7 instances per image, which is significantly more than other prominent datasets. MS COCO also contains images from various perspectives. It also introduced a more stringent method for measuring the detector's efficacy. In contrast to the Pascal VOC and ILSVCR, it calculates the IoU from 0.5 to 0.95 in 0.5-step increments, then employs a combination of these 10 values to determine the Average Precision (AP) metric. In addition, it utilizes AP separately for small, medium, and large objects to compare performance at various scales.

3.2.4 Open Image

Open Images [82] is annotated with image-level annotations, object bounding boxes, and segmentation masks, among others, and contains 9.2 million images. It was introduced in 2017 and has received six enhancements since then. Open Images has 16 million bounding boxes on 1.9 million images for 600 categories, making it the largest dataset for object localization. Its creators took extra care to select images that are fascinating, complex, and diverse, with 8.3 object categories per image. Several modifications were made to the AP introduced in Pascal VOC, such as disregarding unannotated classes, requiring class and subclass detection, etc.

3.2.5 Issues of Data Skew/Bias

A perceptive reader would undoubtedly observe that the number of images for different classes varies substantially across all datasets [86]. Three of the four datasets under discussion (Pascal VOC, MS-COCO, and Open Images Dataset) have a very significant decrease in the number of images beyond the top five most prevalent classes. In MS-COCO dataset, there are 13775 images containing 'person' and then 2829 images containing 'car'. The number of images for the remaining 18 classes in this dataset nearly decreases linearly to 55 for the class'sheep'. Similarly, the class 'person' in the MS-COCO dataset contains 262465 images, while the class 'car' contains 43867 images. The downward trend continues until only 198 images for the class 'hair dryer' remain. The class 'Man' is the most common in the Open Images Dataset, with 378077 images, whereas the class 'Paper Cutter' contains only three. This plainly represents an imbalance in the datasets and is destined to introduce bias into the object detection model training process. Consequently, an object detection model trained on these biased datasets will most likely demonstrate superior detection performance for classes with a greater number of images in the training data. Although still present, this issue is less pronounced in the ImageNet dataset, where the most frequent class, 'koala,' contains 2469 images and the least frequent class, 'cart,' contains 624 images. In the ImageNet dataset, the two most frequent classes are 'koala' and 'computer keyboard', which are clearly not the most desired objects in a real-world object detection scenario (where people, cars, traffic signs, etc. are of greater concern).

RQ3: **What are the research gaps and prospective future research directions related to visual impairment application research?**

3.3 Research gaps and prospective future research directions

3.3.1 Vision Care

In the information age, individuals use their eyeballs significantly more than in previous decades. They depend on video display terminals such as computers and televisions for an increasing number of activities [87]–[89]. In every country, the number of individuals suffering from visual fatigue has increased geometrically [88]. The perception of vision care, however, has not kept pace. There are numerous methods to alleviate eye strain, prevent myopia, and delay the progression of vision loss.

Feng, D., et al. (2022) [90] seeks to defend children's vision during their formative years by conducting research and developing products. The paper [91] analyzes the vision of university students prior to their participation in athletics and exempts three athletes whose vision did not satisfy the requirements. Following this, athletes are provided eye treatment.

Additionally, Weise [92] investigates the impacts of various football helmet visors on athletes. When we are youthful, we must be aware of the need to safeguard our eyesight as soon as feasible. In addition, as we age, the structural and functional integrity of the visual system declines, and the incidence of eye disease increases. During the epidemic, visually impaired individuals who sought medical care faced increased hazards and requirements (167]. Therefore, it is essential to investigate methods to assist the visually impaired.

3.3.2 Smart Home

Even in their own home, the visually impaired cannot move around entirely freely; therefore, a method for controlling smart homes must be developed. The system [93] begins with the smartphone's microphone, Android processes the results, and then connects to NodeMCU via Wi-Fi; however, this system can only manage a limited number of command types. The paper [94] investigates a smart home technology based on touch technology that was originally intended to aid elderly individuals in living independently. However, it could also be used to assist blind individuals live independently in their residences.

The proposed method [95] is also touch-based. Touching the floor plan to receive haptic feedback enables blind individuals to investigate a multi-room floor plan. Literature review revealed that medical aides for the visually impaired and home assistive device technology are still in the developmental stage.

No device has been tested permanently on visually impaired individuals. Even if they all have a visual impairment, their visual conditions and assistive requirements differ considerably. Some individuals require magnification devices, whereas others require portable electronic aides or tools that can read documents audibly. However, the majority of current assistive devices are blind poles or spectacles. The blind stick only enables the visually impaired individual to access information about an obstacle in a limited area in front of them, thereby preventing them from gaining a comprehensive understanding of their surroundings.

Despite the undeniable utility of these aides, they are less diverse and specialized. Large, complicated, and expensive, these aides do not meet the actual requirements of the visually impaired community. These deficiencies place additional strain on visually impaired individuals, who already have difficulty working and limited incomes.

In the future, we anticipate that visually impaired assistive devices will be more intelligent and applicable in a wider variety of situations. These new devices will be more specialized and portable, making the visually impaired's lives significantly more convenient.

4. Conclusion

Even though AI has made significant progress over the past decade, the top detectors are still far from performance saturation. As its real-world applications expand, the demand for lightweight models that can be deployed on mobile and embedded systems will grow exponentially. Wearable devices will be applicable in a greater diversity of everyday situations. Several publications on devices utilising artificial intelligence were examined. These AI-powered devices can help visually impaired people read text, navigate the street, and even create art. The majority of assisted reading research focuses on creating audio for visually impaired individuals to comprehend. The preponderance of AI-powered mobility aides take the form of essentially identical blind canes and glasses. Enhancing their quality of life, visually impaired individuals can create art with the aid of intelligent devices. There is a growing interest in this field, but it remains an unresolved problem. In this paper, we demonstrate how numerous object detectors have evolved since their inception. While two-stage detectors are generally more accurate, they are too sluggish for use in real-time applications such as autonomous vehicles or security. In recent years, however, one-stage detectors have become more accurate and quicker than two-stage detectors. Deeplearning is also certain to shock people. SwinTransformer, a transformer-based detector, is the most accurate detector to date. With the current upward trend in the precision of detectors, we have high aspirations for the development of more precise and quicker detectors.

Appendix A

Table A1. Full Query Term in Publication Databases

Scientific Database	Query	Studies Results #
ACM Digital Library	((deep learning AND ((fft[Filter]) AND (english[Filter]) AND (2020:2023[pdat]))) AND (Visual Impairment OR Blindness OR Visionless AND ((fft[Filter]) AND (english[Filter]) AND (2020:2023[pdat])))) AND ("visual impairment" AND assistive technology AND ((fft[Filter]) AND (english[Filter]) AND (2020:2023[pdat]))) AND ((fft[Filter]) AND (english[Filter])) AND ((fft[Filter]) AND (english[Filter]))	20
IEEE Xplore	("All Metadata":deep learning) AND ("All Metadata":visual impairment OR "All Metadata":blindness OR "All Metadata":visionless) AND ("All Metadata":assistive technology OR "All Metadata":visual impairment)	19
ScienceDirect	(("Visual Impairment") OR (Blindness) OR (visionless")) AND ("deep learning") AND ("Assistive Technology" OR "Visual Impairment")	150
SpringerLink	("deep learning") AND (("visual impairment" OR "blindness" OR "visionless")) AND (("assistive technology") OR ("Visual Impairment"))	102
Scopus	TITLE-ABS-KEY (("deep learning") AND ("visual impairment" OR "blindness" OR "tensionless") AND (("assistive technology") OR ("assistive technology")))	31

References

[1] P. Argüeso, "Human ocular mucins: The endowed guardians of sight," *Adv. Drug Deliv. Rev.*, vol. 180, 2022. https://doi.org/10.1016/j.addr.2021.114074

[2] P. R. Boyce, "Light, lighting and human health," *Light. Res. Technol.*, vol. 54, no. 2, pp. 101–144, 2022. https://doi.org/10.1177/14771535211010267

[3] S. Ferrari *et al.*, "Presence of SARS-CoV-2 RNA in human corneal tissues donated in Italy during the COVID-19 pandemic," *BMJ Open Ophthalmol.*, vol. 7, no. 1, pp. 5–9, 2022. http://dx.doi.org/10.1136/bmjophth-2022-000990

[4] S. G. Singh Gustavo; Shah, Ruchi; Kramerov, Andrei A.; Wright, Robert Emery; Spektor, Tanya M; Ljubimov, Alexander V.; Arumugaswami, Vaithilingaraja; Kumar, Ashok, "SARS-CoV-2 and its beta variant of concern infect human conjunctival epithelial cells and induce differential antiviral innate immune response.," *Ocul. Surf.*, vol. 23, no. NA, pp. 184–194, 2021. https://doi.org/10.1016/j.jtos.2021.09.007

[5] M. M. C. Schwarz Kaleigh A; Davoli, Katherine A; McMillen, Cynthia M; Albe, Joseph R; Hoehl, Ryan M; Demers, Matthew J; Ganaie, Safder S; Price, David A; Leung, Daisy W; Amarasinghe, Gaya K; McElroy, Anita K; Reed, Douglas S; Hartman, Amy L, "Rift Valley Fever Virus Infects the Posterior Segment of the Eye and Induces Inflammation in a Rat Model of Ocular Disease.," *J. Virol.*, vol. 96, no. 20, pp. e0111222-NA, 2022. https://doi.org/10.1128/jvi.01112-22

[6] https://www.who.int/publications/i/item/9789241516570.

[7] B. Kuriakose, R. Shrestha, and F. E. Sandnes, “Tools and technologies for blind and visually impaired navigation support: a review,” *IETE Tech. Rev.*, 2022. https://doi.org/10.1080/02564602.2020.1819893
[8] E. L. Cardillo Changzhi; Caddemi, Alina, “Millimeter-Wave Radar Cane: A Blind People Aid With Moving Human Recognition Capabilities,” *IEEE J. Electromagn. RF Microwaves Med. Biol.*, vol. 6, no. 2, pp. 204–211, 2022. https://doi.org/10.1109/JERM.2021.3117129
[9] A. P. Budrionis Darius; Daniušis, Povilas; Indrulionis, Audrius, “Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review.,” *Assist. Technol.*, vol. 34, no. 2, pp. 1–17, 2020. https://doi.org/10.1080/10400435.2020.1743381
[10] J. M. K. Nesemann Ram Prasad; Byanju, Raghunandan; Poudyal, Bimal; Bhandari, Gopal; Bhandari, Sadhan; O'Brien, Kieran S; Stevens, Valerie M; Melo, Jason S; Keenan, Jeremy D., “Association of visual impairment with disability: a population-based study.,” *Eye (Lond).*, vol. 36, no. 3, pp. 1–7, 2021. https://doi.org/10.1038/s41433-021-01498-x
[11] V. P. Yasin Peniarsih; Gozali, Ahmad; Junaedi, Ifan, “Application of expert system diagnosis of color blindness with ishihara method with microsoft vb 6.0,” *Int. J. Informatics, Econ. Manag. Sci.*, vol. 1, no. 1, p. 13, 2022. https://doi.org/10.52362/ijiems.v1i1.678
[12] H. X. Xinghong Liu; Zhuming, Zhang; Menghan, Xia; Chengze, Li; Tien-Tsin, Wong, “Colorblind-shareable videos by synthesizing temporal-coherent polynomial coefficients,” *ACM Trans. Graph.*, vol. 38, no. 6, pp. 1–12, 2019. https://doi.org/10.1145/3355089.3356534
[13] R. T. Alcaraz Martínez Mireia Ribera; Granollers Saltiveri, Toni, “Methodology for heuristic evaluation of the accessibility of statistical charts for people with low vision and color vision deficiency,” *Univers. Access Inf. Soc.*, vol. 21, no. 4, pp. 863–894, 2021. https://doi.org/10.1007/s10209-021-00816-0
[14] M. H. S. Mahjoob Javad; Anderson, Andrew J, “The effect of mental load on psychophysical and visual evoked potential visual acuity.,” *Ophthalmic Physiol. Opt.*, vol. 42, no. 3, pp. 586–593, 2022. https://doi.org/10.1111/opo.12955
[15] M. H. S. Mahjoob Javad; Anderson, Andrew J, “The effect of mental load on psychophysical and visual evoked potential visual acuity.,” *Ophthalmic Physiol. Opt.*, vol. 42, no. 3, pp. 586–593, 2022. https://doi.org/10.1111/opo.12955
[16] V. C. D. De Cock Pauline; Leu-Semenescu, Smaranda; Aerts, Cécile; Castelnovo, Giovanni; Abril, Beatriz; Drapier, Sophie; Olivet, Hélène; Corbillé, Anne-Gaëlle; Leclair-Visonneau, Laurène; Sallansonnet-Froment, Magali; Lebouteux, Marie; Anheim, Mathieu; Ruppert, E, “Safety and efficacy of subcutaneous night-time only apomorphine infusion to treat insomnia in patients with Parkinson's disease (APOMORPHEE): a multicentre, randomised, controlled, double-blind crossover study.,” *Lancet. Neurol.*, vol. 21, no. 5, pp. 428–437, 2022. https://doi.org/10.1016/s1474-4422(22)00085-0
[17] F. A. Almutairi Nawaf; Ahmad, Khabir; Magliyah, Moustafa S.; Schatz, Patrik, “Congenital stationary night blindness: an update and review of the disease spectrum in Saudi Arabia.,” *Acta Ophthalmol.*, vol. 99, no. 6, pp. 581–591, 2020. https://doi.org/10.1111/aos.14693
[18] B.-J. Cho, “Congenital Stationary Night Blindness,” in *Inherited Retinal Disease*, Singapore: Springer Nature Singapore, 2022, pp. 117–123.
[19] S. Y. Mochida Takeshi; Nomura, Takuhei; Hatake, Ryoma; Ohno-Matsui, Kyoko, “Association between peripheral visual field defects and focal lamina cribrosa defects in highly myopic eyes.,” *Jpn. J. Ophthalmol.*, vol. 66, no. 3, pp. 285–295, 2022. https://doi.org/10.1007/s10384-022-00909-0
[20] L. E. Donaldson Arshia; Sacco, Simone; Micieli, Jonathan A.; Margolin, Edward, “Junctional Scotoma and Patterns of Visual Field Defects Produced by Lesions Involving the Optic Chiasm.,” *J. Neuroophthalmol.*, vol. 42, no. 1, pp. e203–e208, 2021. https://doi.org/10.1097/wno.0000000000001394
[21] V. N. . D. Vakharia Beate; Tisdall, Martin, “Visual field defects in temporal lobe epilepsy surgery.,” *Curr. Opin. Neurol.*, vol. 34, no. 2, pp. 188–196, 2021. https://doi.org/10.1097/wco.0000000000000905
[22] E. M. Ekici Sasan; Hou, Huiyuan; Proudfoot, James A.; Zangwill, Linda M.; L., Jiun; Oh, Won Hyuk; Kamalipour, Alireza; Liebmann, Jeffrey M.; De Moraes, Carlos Gustavo; Girkin, Christopher A.; El-Nimri, Nevin W.; Weinreb, Robert N., “Central Visual Field Defects in Patients with Distinct Glaucomatous Optic Disc Phenotypes,” *Am. J. Ophthalmol.*, vol. 223, no. NA, pp. 229–240, 2020. https://doi.org/10.1016/j.ajo.2020.10.015
[23] B. S. Schmitz Katherine L; Wingerson, Mathew J; Walker, Gregory A; Wilson, Julie C; Howell, David R, “Double Vision and Light Sensitivity Symptoms are Associated With Return-to-School Timing After Pediatric Concussion.,” *Clin. J. Sport Med.*, vol. Publish Ah, 2022. https://doi.org/10.1097/jsm.0000000000001106
[24] R. K. Bartlett Yi Xuan; Hourcade, Juan Pablo; Rector, Kyle, “Exploring the Opportunities for Technologies to Enhance Quality of Life with People who have Experienced Vision Loss,” *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, p. 191, 2019. https://doi.org/10.1145/3290605.3300421
[25] S. W. Bi Cong; Zhang, Jilong; Huang, Wutao; Wu, Bochun; Gong, Yi; Ni, Wei, “A Survey on Artificial Intelligence Aided Internet-of-Things Technologies in Emerging Smart Libraries.,” *Sensors (Basel).*, vol. 22, no. 8, p. 2991, 2022. https://doi.org/10.3390/s22082991
[26] L. Z. Jia Zhi; Xu, Fei; Jin, Hai, “Cost-Efficient Continuous Edge Learning for Artificial Intelligence of Things,” *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7325–7337, 2022. https://doi.org/10.1109/JIOT.2021.3104089
[27] S. O. Zhu Kaoru; Dong, Mianxiong, “Energy-Efficient Artificial Intelligence of Things With Intelligent Edge,” *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7525–7532, 2022. https://doi.org/10.1109/JIOT.2022.3143722
[28] D. N. A. Nya Hassane, “Model-Free Control Policies for Inventory Management in Supply Chain,” *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*, vol. NA, no. NA. p. NA-NA, 2022. https://doi.org/10.1109/CoDIT55151.2022.9803914
[29] Y. B. Mashayekhy Amir; Yuan, Xue-Ming; Xue, Anrong, “Impact of Internet of Things (IoT) on Inventory Management: A Literature Survey,” *Logistics*, vol. 6, no. 2, p. 33, 2022. https://doi.org/10.3390/logistics6020033
[30] Y.-M. H. Tang George To Sum; Lau, Yui-Yip; Tsui, Shuk-Ying, “Integrated Smart Warehouse and Manufacturing Management with Demand Forecasting in Small-Scale Cyclical Industries,” *Machines*, vol. 10, no. 6, p. 472, 2022. https://doi.org/10.3390/machines10060472
[31] M. K. Mahobe Pradeep; Jha, Shashi Shekhar, “Nature-Inspired AI Techniques in Intelligent Transportation System,” in *Lecture Notes in Electrical Engineering*, Springer Nature Singapore, 2022, pp. 251–263. https://doi.org/10.1007/978-981-16-8542-2_20
[32] D. K. Loske Matthias, “Human-AI collaboration in route planning: An empirical efficiency-based analysis in retail logistics,” *Int. J. Prod. Econ.*, vol. 241, no. NA, pp. 108236-NA, 2021. https://doi.org/10.1016/j.ijpe.2021.108236
[33] I. Z. Lee Helen; Moore, Kate; Zhou, Xiaofei; Perret, Beatriz; Cheng, Yihong; Zheng, Ruiying; Pu, Grace, “AI Book Club,” *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education*. ACM, 2022. https://doi.org/10.1145/3478431.3499318
[34] T. P. Kabudi Ilias O.; Olsen, Dag H., “AI-enabled adaptive learning systems: A systematic mapping of the literature,” *Comput. Educ. Artif. Intell.*, vol. 2, p. 100017, 2021. https://doi.org/10.1016/j.caeai.2021.100017
[35] Y. A. Alsaawy Ahmad; Abi Sen, Adnan; Alshanqiti, Abdullah; Bhat, Wasim Ahmad; Bahbouh, Nour Mahmoud, “A Comprehensive and Effective Framework for Traffic Congestion Problem Based on the Integration of IoT and Data Analytics,” *Appl. Sci.*, vol. 12, no. 4, p. 2043, 2022. https://doi.org/10.3390/app12042043
[36] S. P. Bhupathi Nithish, Ankit Panda, Trishala Reddy, Vishwa Gohil, Ishita Kundaliya, “Self-Driving Car to Drive Autonomously using Image Processing and Deep Learning,” *Irjet*, vol. 9, no. 1, pp. 125–132, 2022. http://dx.doi.org/10.13140/RG.2.2.16212.88963/1
[37] S. K. B. M.; Gayathri, S.; Srinidhi, S.; Hemasundari, H.; Sowmiya, S.; Shavan Kumar, S., “AI-Based Motorized Appearance Acknowledgement Scheme for an Attendance Marking System,” in *Advances in Social Networking and Online Communities*, vol. NA, no. NA, 2022, pp. 98–109.

[38] M. Surve, P. Joshi, S. Jamadar, and M. M. N. Vharkate, “Automatic Attendance System using Face Recognition Technique,” *Int. J. Recent Technol. Eng.*, vol. 9, no. 1, pp. 2134–2138, 2020. http://www.doi.org/10.35940/ijrte.A2644.059120
[39] Y. Himeur *et al.*, *AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives*, no. 0123456789. Springer Netherlands, 2022.
[40] A. Y. M. Chakor Azmani; Abdellah, Azmani, “Proposing a Layer to Integrate the Sub-classification of Monitoring Operations Based on AI and Big Data to Improve Efficiency of Information Technology Supervision,” *Appl. Comput. Syst.*, vol. 27, no. 1, pp. 43–54, 2022. https://doi.org/10.2478/acss-2022-0005
[41] S. A. Upadhyaya Anushri; Rengaraj, Venkatesh; Srinivasan, Kavitha; Casey, Paula Anne Newman; Schehlein, Emily M., “Validation of a portable, non-mydriatic fundus camera compared to gold standard dilated fundus examination using slit lamp biomicroscopy for assessing the optic disc for glaucoma.,” *Eye (Lond).*, vol. 36, no. 2, pp. 1–7, 2021. https://doi.org/10.1038/s41433-021-01485-2
[42] D. Mishra, S. Gade, K. Glover, R. Sheshala, and T. R. R. Singh, “Vitreous Humor: Composition, Characteristics and Implication on Intravitreal Drug Delivery,” *Curr. Eye Res.*, vol. 48, no. 2, pp. 208–218, 2022. https://doi.org/10.1080/02713683.2022.2119254
[43] R. F. V. Spaide Philippe; Maloca, Peter M; Scholl, Hendrik P N; Otto, Tilman P; Caujolle, Sophie, “Imaging The Vitreous With A Novel Boosted Optical Coherence Tomography Technique: Vitreous Degeneration and Cisterns.,” *Retina*, vol. 42, no. 8, pp. 1433–1441, 2022. https://doi.org/10.1097/iae.0000000000003474.
[44] S. M. . R. Zekavat Vineet K.; Trinder, Mark; Ye, Yixuan; Koyama, Satoshi; Honigberg, Michael C.; Yu, Zhi; Pampana, Akhil; Urbut, Sarah; Haidermota, Sara; O’Regan, Declan P.; Zhao, Hongyu; Ellinor, Patrick T.; Segrè, Ayellet V.; Elze, Tobias; Wiggs, Janey L.; Marton, “Deep Learning of the Retina Enables Phenome- and Genome-wide Analyses of the Microvasculature.,” *Circulation*, vol. 145, no. 2, pp. 134–150, 2021. https://doi.org/10.1161/CIRCULATIONAHA.121.057709
[45] T. M. Izumi Ichiro; Kawano, Taizo; Sakaihara, Manabu; Iida, Tomohiro, “Morphological differences of choroid in central serous chorioretinopathy determined by ultra-widefield optical coherence tomography.,” *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 260, no. 1, pp. 1–7, 2021. https://doi.org/10.1007/s00417-021-05380-0
[46] X. Xu *et al.*, “Automatic Segmentation and Measurement of Choroid Layer in High Myopia for OCT Imaging Using Deep Learning,” *J. Digit. Imaging*, vol. 35, no. 5, pp. 1153–1163, Oct. 2022. https://doi.org/10.1007/s10278-021-00571-x
[47] B. G.-V. Burgos-Blasco Noemi; Vidal-Villegas, Beatriz; Martinez-de-la-Casa, Jose M.; Donate-Lopez, Juan; Martín-Sánchez, Francisco Javier; González-Armengol, Juan Jorge; Porta-Etessam, Jesús; Martin, José Luis R.; Garcia-Feijoo, Julian, “Optic nerve and macular optical coherence tomography in recovered COVID-19 patients.,” *Eur. J. Ophthalmol.*, vol. 32, no. 1, pp. 11206721211001020–11206721211001020, 2021. https://doi.org/10.1177/11206721211001019
[48] V. D.-M. Biousse Helen V; Saindane, Amit M; Lamirel, Cédric; Newman, Nancy J, “Imaging of the optic nerve: technological advances and future prospects.,” *Lancet. Neurol.*, vol. 21, no. 12, pp. 1135–1150, 2022. https://doi.org/10.1016/s1474-4422(22)00173-9
[49] M. J. . M. Page Joanne E.; Bossuyt, Patrick M.M.; Boutron, Isabelle; Hoffmann, Tammy; Mulrow, Cynthia D.; Shamseer, Larissa; Tetzlaff, Jennifer; Akl, Elie A.; Brennan, Sue E.; Chou, Roger; Glanville, Julie; Grimshaw, Jeremy M.; Hróbjartsson, Asbjørn; Lalu, Manoj M., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” *BMJ*, vol. 372, p. n71, 2021. https://doi.org/10.1136/bmj.n71
[50] PRISMA Endorsers, “PRISMA.”.
[51] M. McDonagh, K. Peterson, P. Raina, S. Chang, and P. Shekelle, “Avoiding Bias in Selecting Studies.,” Rockville (MD), 2008.
[52] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, “A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using RNA-seq data,” *Comput. Methods Programs Biomed.*, vol. 166, pp. 99–105, 2018. https://doi.org/10.1016/j.cmpb.2018.10.004
[53] R. Tapu, B. Mocanu, and T. Zaharia, “DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance,” *Sensors (Switzerland)*, vol. 17, no. 11, 2017. https://doi.org/10.3390/s17112473
[54] A. Nagarajan and G. M P, “Hybrid Optimization-Enabled Deep Learning for Indoor Object Detection and Distance Estimation to Assist Visually Impaired Persons,” *Adv. Eng. Softw.*, vol. 176, no. July 2022, p. 103362, 2023. https://doi.org/10.1016/j.advengsoft.2022.103362
[55] J. Ganesan, A. T. Azar, S. Alsenan, N. A. Kamal, B. Qureshi, and A. E. Hassanien, “Deep Learning Reader for Visually Impaired,” *Electron.*, vol. 11, no. 20, 2022. https://doi.org/10.3390/electronics11203335
[56] M. M. Islam, S. Nooruddin, F. Karray, and G. Muhammad, “Multi-level feature fusion for multimodal human activity recognition in Internet of Healthcare Things,” *Inf. Fusion*, vol. 94, pp. 17–31, 2023. https://doi.org/10.1016/j.inffus.2023.01.015
[57] M. Zounemat-Kermani and A. Mahdavi-Meymand, “Hybrid meta-heuristics artificial intelligence models in simulating discharge passing the piano key weirs,” *J. Hydrol.*, 2019. https://doi.org/10.1016/j.jhydrol.2018.11.052
[58] R. O. M. Ogundokun Rytis; Damaševičius, Robertas, “Human Posture Detection Using Image Augmentation and Hyperparameter-Optimized Transfer Learning Algorithms,” *Appl. Sci.*, vol. 12, no. 19, p. 10156, 2022. https://doi.org/10.3390/app121910156
[59] P. F. . G. Felzenszwalb Ross; McAllester, David; Ramanan, Deva, “Object Detection with Discriminatively Trained Part-Based Models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, 2010. https://doi.org/10.1109/TPAMI.2009.167
[60] M. Ali, F. Sahin, S. Kumar, and C. Savur, “360° view camera based visual assistive technology for contextual scene information,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017*, 2017, vol. 2017-Janua, pp. 2135–2140. https://doi.org/10.1109/SMC.2017.8122935
[61] Y. B. LeCun Bernhard E.; Denker, John S.; Henderson, D.; Howard, Richard; Hubbard, W.; Jackel, Lawrence D., “Backpropagation applied to handwritten zip code recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989. https://doi.org/10.1162/neco.1989.1.4.541
[62] C. Playout, R. Duval, M. C. Boucher, and F. Cheriet, “Focused Attention in Transformers for interpretable classification of retinal images,” *Med. Image Anal.*, vol. 82, no. July, 2022. https://doi.org/10.1016/j.media.2022.102608
[63] R. Fan *et al.*, “Detecting Glaucoma from Fundus Photographs Using Deep Learning without Convolutions: Transformer for Improved Generalization,” *Ophthalmol. Sci.*, vol. 3, no. 1, p. 100233, 2023. https://doi.org/10.1016/j.xops.2022.100233
[64] F. S. Shi Wei; Duan, Huiyu; Liu, Xiaotian; Hu, Menghan; Wang, Wei; Zhai, Guangtao, “Drawing reveals hallmarks of children with autism,” *Displays*, vol. 67, no. NA, pp. 102000-NA, 2021. https://doi.org/10.1016/j.displa.2021.102000
[65] G. S. R. . D. Satyanarayana Prashant; Das, Santos Kumar, “Vehicle detection and classification with spatio-temporal information obtained from CNN,” *Displays*, vol. 75, no. NA, p. 102294, 2022. https://doi.org/10.1016/j.displa.2022.102294
[66] C. Z. Hua Baojiang; Song, Weigang; Yang, Jianyu, “Circular coding: A technique for visual localization in urban areas,” *Displays*, vol. 75, no. NA, p. 102299, 2022. https://doi.org/10.1016/j.displa.2022.102299
[67] S. S. Feuerriegel Yash Raj; von Krogh, Georg; Zhang, Ce, “Bringing artificial intelligence to business management,” *Nat. Mach. Intell.*, vol. 4, no. 7, pp. 611–613, 2022. https://doi.org/10.1038/s42256-022-00512-5
[68] X. W. Wang Chen; Liu, Bing; Zhou, Xiaoqing; Zhang, Liang; Zheng, Jin; Bai, Xiao, “Multi-view stereo in the Deep Learning Era: A comprehensive revfiew,” *Displays*, vol. 70, no. NA, pp. 102102-NA, 2021. https://doi.org/10.1016/j.displa.2021.102102
[69] L. L. Mohammadpour Teck Chaw; Liew, Chee Sun; Aryanfar, Alihossein, “A Survey of CNN-Based Network Intrusion Detection,” *Appl. Sci.*, vol. 12, no. 16, p. 8162, 2022. https://doi.org/10.3390/app12168162

[70] M. Mukhiddinov, A. B. Abdusalomov, and J. Cho, “Automatic Fire Detection and Notification System Based on Improved YOLOv4 for the Blind and Visually Impaired,” *Sensors*, vol. 22, no. 9, 2022. https://doi.org/10.3390/s22093307

[71] A. Mogadala, M. Kalimuthu, and D. Klakow, “Trends in integration of vision and language research: A survey of tasks, datasets, and methods,” *J. Artif. Intell. …*, 2021.

[72] H. S. Phillips Shelly; Klang, Eyal, “Oncological Applications of Deep Learning Generative Adversarial Networks.,” *JAMA Oncol.*, vol. 8, no. 5, pp. 677-NA, 2022. https://doi.org/10.1001/jamaoncol.2021.8202

[73] S. Shahriar, “GAN computers generate arts? A survey on visual arts, music, and literary text generation using generative adversarial network,” *Displays*, vol. 73, no. NA, pp. 102237-NA, 2022. https://doi.org/10.1016/j.displa.2022.102237

[74] A. Iqbal, M. Sharif, M. Yasmin, M. Raza, and S. Aftab, “Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey,” *Int. J. Multimed. Inf. Retr.*, vol. 11, no. 3, pp. 333–368, 2022. https://doi.org/10.1007/s13735-022-00240-x

[75] C. Qian, J. Zhu, Y. Shen, Q. Jiang, and Q. Zhang, “Deep Transfer Learning in Mechanical Intelligent Fault Diagnosis: Application and Challenge,” *Neural Process. Lett.*, vol. 54, no. 3, pp. 2509–2531, 2022. https://doi.org/10.1007/s11063-021-10719-z

[76] R. O. Ogundokun, R. Maskeliūnas, and R. Damaševičius, “Human Posture Detection Using Image Augmentation and Hyperparameter-Optimized Transfer Learning Algorithms,” *Applied Sciences*, vol. 12, no. 19. 2022. https://doi.org/10.3390/app121910156

[77] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.” 2016.

[78] Y. Ganin *et al.*, “Domain-Adversarial Training of Neural Networks.” 2016.

[79] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The Pascal Visual Object Classes (VOC) Challenge,” *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010. https://doi.org/10.1007/s11263-009-0275-4

[80] O. D. Russakovsky Jia; Su, Hao; Krause, Jonathan; Satheesh, Sanjeev; Ma, Sean; Huang, Zhiheng; Karpathy, Andrej; Khosla, Aditya; Bernstein, Michael S.; Berg, Alexander C.; Fei-Fei, Li, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015. https://doi.org/10.1007/s11263-015-0816-y

[81] C. L. T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár and Zitnick, “Microsoft COCO: common objects in context,” in *D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), Computer Vision – ECCV 2014*, 2014, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

[82] A. Kuznetsova *et al.*, “The Open Images Dataset V4,” *Int. J. Comput. Vis.*, vol. 128, no. 7, pp. 1956–1981, Jul. 2020. https://doi.org/10.1007/s11263-020-01316-z

[83] A. Z. M. Everingham, L. Van Gool, C.K.I. Williams, J. Winn, “The PASCAL Visual Object Classes Challenge 2007 (VOC2007).” .

[84] J. W. M. Everingham, “The PASCAL visual object classes challenge 2012 (VOC2012) development kit 32.” .

[85] L. F.-F. J. Deng, W. Dong, R. Socher, L. Li, Kai Li, “ImageNet: a large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. https://doi.org/10.1109/CVPR.2009.5206848

[86] A. Aslam and E. Curry, “A Survey on Object Detection for the Internet of Multimedia Things (IoMT) using Deep Learning and Event-based Middleware: Approaches, Challenges, and Future Directions,” *Image Vis. Comput.*, vol. 106, p. 104095, Feb. 2021. https://doi.org/10.1016/j.imavis.2020.104095

[87] N. E. Ongalia Titiek; Meryana, Pauline, “COMPUTER VISION SYNDROME IN MEDICAL STUDENTS IN THE ERA OF THE COVID-19 PANDEMIC,” *J. Widya Med. Jr.*, vol. 4, no. 5, pp. 199–204, 2022. https://doi.org/10.33508/jwmj.v4i3.4096

[88] H. Ö. Öztürk Bediz, “The Effects of Smartphone, Tablet and Computer Overuse on Children’s Eyes During the COVID-19 Pandemic,” *J. Pediatr. Res.*, vol. 8, no. 4, pp. 491–497, 2021. https://doi.org/10.4274/jpr.galenos.2021.72623

[89] M. M. A. Zalat Soliman; Wassif, Ghada A.; Tarhouny, Shereen A. El; Mansour, Tayseer M., “Computer vision syndrome, visual ergonomics and amelioration among staff members in a Saudi medical college,” *Int. J. Occup. Saf. Ergon.*, vol. 28, no. 2, pp. 1–9, 2021. https://doi.org/10.1080/10803548.2021.1877928

[90] D. Feng, C. Lu, Q. Cai, and J. Lu, “A Study on the Design of Vision Protection Products Based on Children’s Visual Fatigue under Online Learning Scenarios,” *Healthcare*, vol. 10, no. 4, p. 621, Mar. 2022. https://doi.org/10.3390/healthcare10040621

[91] K. K. . G. Weise Sarah J.; Hale, M Heath; Springer, Daniel B.; Swanson, Mark W., “Pre-participation Vision Screening and Comprehensive Eye Care in National Collegiate Athletic Association Athletes.,” *Optom. Vis. Sci.*, vol. 98, no. 7, pp. 764–770, 2021. https://doi.org/10.1097/opx.0000000000001738

[92] K. K. . S. Weise Mark W.; Galt, Sarah J.; Springer, Daniel B.; Crosson, Jason N.; DeCarlo, Dawn K.; Hale, Matthew Heath; Nicholson, Joshua Ryne; Robinson, James B., “Objective Vision-related Indications for Clear and Tinted Football Helmet Visors.,” *Optom. Vis. Sci.*, vol. 98, no. 7, pp. 833–838, 2021. https://doi.org/10.1097/opx.0000000000001730

[93] L. Y. Triyono Tri Raharjo; Sukamto, Sukamto; Hestinigsih, I, “VeRO: Smart home assistant for blind with voice recognition,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1108, no. 1, pp. 012016-NA, 2021. https://doi.org/10.1088/1757-899x/1108/1/012016

[94] N. O. Yu Ziwei; Wang, Hehe; Tao, Da; Jing, Liang, “The Effects of Smart Home Interface Touch Button Design Features on Performance among Young and Senior Users.,” *Int. J. Environ. Res. Public Health*, vol. 19, no. 4, p. 2391, 2022. https://doi.org/10.3390/ijerph19042391

[95] S. K. Feitl Julian; Götzelmann, Timo, “Accessible Electrostatic Surface Haptics: Towards an Interactive Audiotactile Map Interface for People With Visual Impairments,” *The15th International Conference on PErvasive Technologies Related to Assistive Environments*, vol. NA, no. NA. p. NA-NA, 2022. https://doi.org/10.1145/3529190.3534781