Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
15/1 (2022), 21-31. DOI: http://dx.doi.org/10.21609/jiki.v15i1.1017

A Systematic Literature Review on SOTA Machine learning-supported Computer Vision Approaches to Image Enhancement

Marco Klaiber[1], Jonas Klopfer[2]

[1,2]Aalen University, Research Center of Data Science, Beethovenstraße 1, 73430 Aalen, Germany

Email:marco.klaiber@studmail.htw-aalen.de[1], jonas.klopfer@studmail.htw-aalen.de[2]

Abstract

Image enhancement as a problem-oriented process of optimizing visual appearances to provide easier-to-process input to automated image processing techniques is an area that will consistently be a companion to computer vision despite advances in image acquisition and its relevance continues to grow. For our systematic literature review, we consider the major peer-reviewed journals and conference papers on the state of the art in machine learning-based computer vision approaches for image enhancement. We describe the image enhancement methods relevant to our work and introduce the machine learning models used. We then provide a comprehensive overview of the different application areas and formulate research gaps for future scientific work on machine learning based computer vision approaches for image enhancement based on our results.

Keywords: *Image enhancement, Machine learning, Computer vision, ML, CV*

1. Introduction

Image enhancement is a problem-oriented process with the goal of optimizing visual appearance or providing a "better" – i.e., easier to process – input to automated image processing techniques [1–3]. Improving input images is necessary because downstream models rely on the processed input to perform their task qualitatively and as expected, which is of top priority in a wide variety of fields such as medicine, military, and autonomous driving [4–6]. In such automated image processing techniques, so-called machine learning (ML) is usually used. ML is classified under the generic term Artificial Intelligence (AI) and is used to describe this artificial intelligence, which is the "intelligence" of a machine that does not possess any natural intelligence [7, 8]. In this context, research also talks about the fact that there can be no intelligence without learning [7]. ML is defined by being a distinct, evolving branch of computer algorithms that enable human intelligence to learn from the environment, often achieved through computer vision [9]. Computer vision is a method that uses computers to gain a comprehensive understanding of the content or meaning of individual images or videos from digital, i.e., visual, input data [10, 11]. The information obtained is used, for example, for image classification or face recognition [10]. Without image enhancement capabilities, the potential in different application areas of ML would not be exploited because the input data would not be of the required quality, which is why a comprehensive overview of current techniques for image enhancement is relevant to provide a basis for future work dealing with this topic, which should provide sound knowledge in a compact way and enable innovation.

The following paper deals with the current state of research on ML-based computer vision approaches for image enhancement. It shows what they are needed for and which methods are used for this purpose. The individual image enhancement methods relevant to this work are pointed out and described in detail. Subsequently, the investigated ML models and their manifold applications with respect to image enhancement and computer vision are explained and put into a super-ordinate context.

2. Method

To filter out the relevant results from the published literature, a systematic literature search with the search word combination ”computer vision” and ”image enhancement” was conducted in the literature databases IEEEXplore DL and ScienceDirect. The search yielded 217 articles that contained the named keyword combination in the abstract or title, or in the abstract or keywords, sorted chronologically by the respective database. This was supplemented by a forward and backward search. To ensure that only significant scientific papers are included in this search, only peer-reviewed journals were considered. To obtain further information about the individual ML models, a systematic literature search with the search word combination ”machine learning” and ”image enhancement” was conducted in the literature databases IEEEXplore DL and ScienceDirect. The search resulted in 73 articles containing the mentioned keyword combination in the abstract or title, respectively, or in the abstract or keywords, again specifically in the respective database. or keywords. This was also supplemented by a forward and backward search. To ensure that only significant scientific papers are included in this research are considered, only peer-reviewed journals and conference papers are being considered. After these 290 articles were manually reviewed using the four-eyes principle, 58 articles were found to be relevant to the subject of the study as defined at the outset. Within our manual review, it was important to us to ensure the directness to machine learning on the one hand, as well as to pay attention to the topicality of the contributions to represent the current state of research as concretely as possible. Furthermore, in addition to fundamental work, innovative work should also be included, whereby consideration was given to relevance. The relevance of the work takes into account the number of citations, which can be an indicator of relevance but does not directly reflect the quality of the work, but which was reviewed and ensured as part of this literature review [12, 13]. The literature search process is illustrated in detail in Fig. 1.

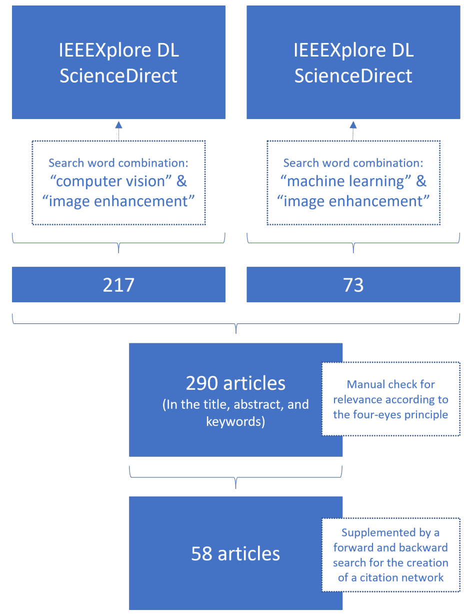

Figure 1. Summary of the search methodology

3. Results

The analyzed ML-supported computer vision approaches to image enhancement have in common that they mostly build on proven image enhancement methods and complement them with ML models. Image enhancement methods are relevant for ML models because they are supposed to provide the scenes ”better” for input into automated image processing models [1, 2]. For example, the image enhancement methods remove noise in scenes [14, 15] and thus facilitate object recognition [16] or improve weak contrasts, dull details, and color variations [17]. In the literature, there are different approaches that categorize image enhancement methods, distinguishing in Histogram Equalization, Retinex, the Dehazing Model, and Deep Learning (DL). [18] and [19] on the other hand share the approaches in Global Enhancement – consisting for example of Linear Amplifying and Histogram Equalization – and Local Enhancement – among others consisting of Deep Neural Networks and Retinex. [20] speaks however of direct and indirect methods. [21] distinguishes between model-based, fusion-based, and learning-based approaches. [19] and [2] divide into the spatial domain Approach - which stands for the direct manipulation of the pixels – and the Frequency Domain method – which makes use of the Fourier transform, for example.

We propose another classification of image enhancement methods, which is visualized in Fig. 3: In our work we distinguish between the classical image enhancement methods - such as Retinex and Histogram Equalization - and the ML methods for image enhancement - such as Random Forest or Support Vector Machines.

3.1. Classical Image Enhancement Methods

Through our research, the Retinex method [17, 18, 22–24], Histogram Equalization [18, 19, 22, 24], and Gaussian Mixture Models (GMM) [16, 25–27] have emerged as the most relevant classical image enhancement methods, which are shown in Fig. 2

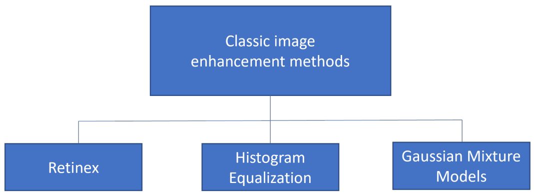

Figure 2. Classification of the classical image enhancement methods

3.1.1. Retinex process. Images or scenes can be affected by atmospheric particles under unclear or foggy conditions [24]. This reduces contrast, changes color, makes objects harder to identify, and makes the image content difficult or impossible to see for the human eye or computer vision [23, 24]. Image dehazing is therefore an important issue for computer vision and this is where the Retinex process can help [24]. Originally, the Retinex process was designed to simulate the human visual system [17]. Today, Retinex-based algorithms are widely used in image enhancement have been adapted and developed further and are now mainly used for applied to reduce and if possible, remove shadows or fog [23, 24]. The Retinex process has a special property; it assumes that the visible light in scenes can be divided into two parts: In the exposure and the reflection [18, 22]. This allows the use of the Retinex process; it uses the reflection properties of objects under the influence of light to extract the required information from the original image and make it visible [23, 24]. The Retinex process is divided into three possible forms: The Single-Scale Retinex (SSR) [14, 17, 22, 24], the Multi-Scale Retinex (MSR) [22, 22, 24, 28] and the Multi-Scale Retinex with Color Restoration (MSRCR) [14, 23]. The SSR works with estimations, which are of Gaussian filters to determine the ambient brightness [22, 24]. The MSR, on the other hand, extends the SSR by several scales to include the equilibrium between color consistency and dynamic range compression [24]. The MSRCR, using Gaussian ambient filters, estimates the exposure of the input scene at various scales and turns up, which is followed by color restoration [23]. The three possible forms of the Retinex process achieve, depending on the input image, a different optimal result. Thus, it cannot be clearly determined whether a of the above-mentioned methods is the ”best” image enhancement method [22–24]. In summary, the Retinex procedure increases the contrast and brightness of the image, improves color consistency, and can work out a dynamic compression [22–24]. In addition, another advantage of the Retinex method is that it can be derived physically, which clearly distinguishes it from subsequent methods such as Historgram Equalization [18].

3.1.2. Histogram Equalization (HE). Histogram Equalization (HE) is the change of the grayscale histogram of the original image from a relative grayscale interval to a uniform distribution over the entire grayscale spectrum [19, 24]. Each pixel receives a grayscale value in the domain [0, 1] [22]. The goal of HE is to maximize the amount of information in an image by improving contrast [19, 24]. Therefore, HE is applied to images with low contrast, e.g. caused by fog or haze. On a technical level, the reduced contrast is achieved by using only a narrow domain of gray scales [24]. Classical HE can be classified as global enhancement because the histogram information of the entire image is used as the basis for enhancement [18, 19]. Here it can be used among other things in domains with a low contrast – such as the background – can lead to so-called over-enhancement, which in turn causes a loss of information [18]. In addition, the number of grayscales is reduced, because pixels are lost which are less frequent in the grayscale distribution. This also leads that details disappear and a loss of information occurs [19]. Some solutions have been proposed for these problems, most of which focus on local enhancement [18, 19, 24]. In [29], [30] and [31] a small window is used, which is displayed pixel by pixel moves through an image and performs a local HE. However, this leads to a significantly increased computational complexity [18, 24]. Also, in [32] an adaptive HE is performed for each individual region. Bilinear interpolation is then used to remove artifacts associated with the edges of each region [19, 32]. In [33], after the local block overlap HE, a filter is inserted for contrast enhancement and noise reduction. Other HE methods, which are particularly suitable for maintaining brightness are found in [20] and [34]. In conclusion, the global HE is more efficient, while the local HE is characterized by higher performance [18, 24].

3.1.3. Gaussian Mixture Model (GMM). The Gaussian mixture model is usually applied at the structural level to express image details and contrasts in scenes more clearly and to obtain a uniform grayscale distribution [16, 25, 26]. GMM is based on probabilities and uses parameters for the normal distribution with deviating mean values, standard deviations, and weightings [25]. Here, the GMM models the individual pixel values of the input scene

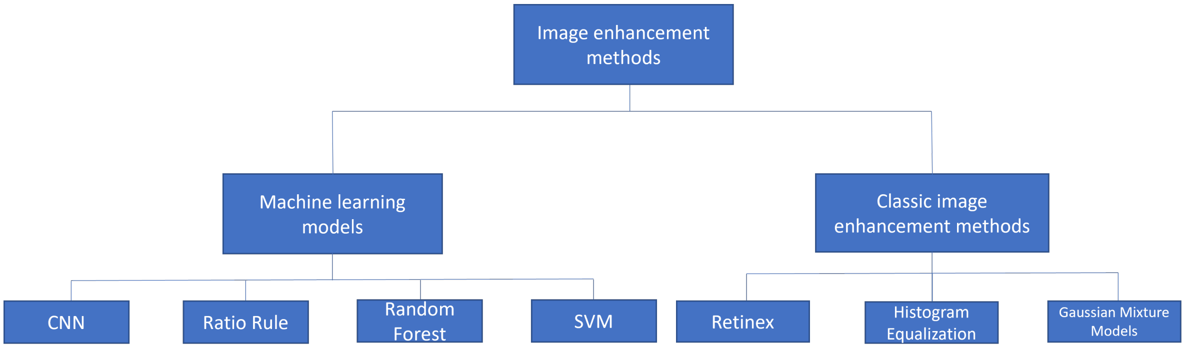

Figure 3. Proposed classification of image enhancement methods

and assigns a weight to each [25]. A component with low standard deviation and high weighting is called ”compact date” and has a density distribution around the mean value [25]. If the standard deviation is larger, the data are distributed over their mean value [25]. Humans themselves are sensitive to widely scattered fluctuations but does not respond to small fluctuations in the compact data [25]. This further scatters data with low standard deviation and in turn compresses data with high scatter, increasing contrast while preserving relevant image details with the prominent structural regions [16, 25]. However, the main disadvantage of GMM is that the signs of the parameters generated using this distribution function are random, since the Gaussian value is symmetrical about zero, which can have a negative effect on the resulting image quality [27].

3.2. Machine Learning Models

As mentioned above, ML is defined by the fact that human intelligence is to be simulated by learning from the environment [9]. At domain of ML based computer vision approaches to image enhancement is to support ML through classical image enhancement methods – such as Retinex or HE [35]. ML models should learn independently, thus without the interaction of a human, and select the optimal filters and thresholds for image enhancement to achieve an optimal result [15, 28, 36]. Moreover, the task of ML is to recognize various patterns, characters, and objects, such as faces, within the scenes to satisfy the previously specified requirements [1, 9, 15, 35, 37]. The different ML models all need to be trained to enable independent learning and progress [8]. There are three basic methods of training: supervised learning, unsupervised learning, and reinforcement learning [8, 15, 35]. In supervised learning, the training set, which the algorithm uses to learn in the training phase, contains labels (inputs and outputs). This means that the model learns which input belongs in which class, receives feedback after each training unit on the extent to which its result is satisfactory and is particularly suitable for classification tasks after a test phase. This means that the algorithm can recognize faces in the application phase [8, 35, 38, 39]. Unsupervised learning, on the other hand, requires it to search for patterns in the inputs itself. The training phase is therefore independent, without labels. An advantage of this method is that the algorithm usually recognizes patterns that humans would not notice, which is in the domain of clustering [8, 15]. Reinforcement learning is based on the carrot and stick principle. After each training session, depending on the quality of the result, either a reward is given, or a penalty is imposed [8]. The advantage of this variant is that the algorithm improves on its own as it strives for the reward [8]. In our extensive research in the domain of image enhancement, the following ML models turned out to be most relevant: The Convolutional Neural Network (CNN) [14, 15, 37, 40], the Ratio Rule (as an example of a Fast Converging Neural Network) [28, 41], the Random Forests [35, 42] and the Support Vector Machine (SVM) model [35, 38, 43], which are all shown in Fig. 4.

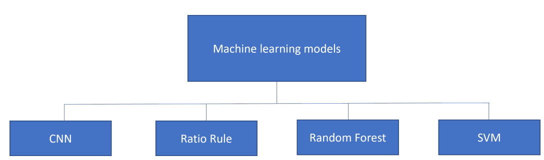

Figure 4. Classification of ML models

3.2.1. Convolutional Neural Network (CNN).

Convolutional Neural Networks (CNNs) belong to the category of methods based on (deep) learning and are effective solutions for classification and recognition problems with large amounts of data

[21, 37, 44]. A CNN has the possibility to display an input scene in a matrix (width x height x color channels). It usually consists of an input and output layer, convolution layers with several additional hidden layers, which are composed of several convolution cores, as well as the pooling layers and a fully connected layer [14, 15, 37, 40, 44–47]. The input scene is primarily analyzed using set filter parameters. The filter moves through the image with a fixed pixel size, detects the structure it was specialized to detect in the training phase and passes the result to the next layer, which in turn detects more complex structures [37, 45, 47]. The hidden layers are mainly used for the detection of structures within the image, e.g. for edge detection [37]. The filter has a fixed weight for each point in its viewport and computes a result matrix from the pixel values in the current viewport and their weights, which is further reduced by the pooling layers [46, 47]. By pooling, only the ”largest” values are passed on, reducing the parameters in the matrix before they are passed on to the next layer [46, 47]. Finally, the object information is passed on to the corresponding class using neurons and the object is recognized and learned by the given characteristics of the class [37, 47]. The training of conventional CNNs requires considerable computing effort and large amounts of data to optimize the weights and biases [37, 44]. For this reason, [37] propose a new fast learning algorithm for CNNs based on an Extreme Learning Machine (ELM). With this method, the weightings can be calculated in a single iteration, which drastically reduces the learning time and enables accurate results with minimal training data [37].

3.2.2. Ratio Rule. The Ratio Rule is a learning algorithm for fast converging neural networks, which offers a solution for the restoration of natural colors of a scene after a grayscale image enhancement – for example with the already described HE – has been performed [3, 28, 41, 48, 49]. This example makes it clear that ML methods often represent an extension of the classical methods or build on them. Grayscale image enhancement within color scenes may not be able to maintain the color relationships between the RGB channels (red, green, blue components), resulting in bias of the color information after enhancement [3, 28]. Such bias occurs when the RGB components are not properly combined [3]. This is exactly where the Ratio Rule comes in: The neural network is used as associative memory to retrieve the natural color characteristics of the image pixels after grayscale enhancement [3, 48]. The algorithm of the Ratio Rule learns by the color relationships of each pixel as a ”line of attraction” [3, 48]. The Ratio Rule creates a line which represents all stable points in the state space in a family pattern [3, 48]. All points with deviations due to external influences approach the ”line of attraction”, allowing the neural network to recover the color information [3, 48]. This ML method is mainly used in domains where color scenes are used. However, the wide range of applications of this method also includes the processing and improvement of old manuscripts that have faded and become hardly legible over time. Here, the Ratio Rule can prepare the texts in such a way that they are again readable and preserved for posterity [3, 28, 41, 48, 49].

3.2.3. Random Forest (RF). The random forest is an ensemble learning method with a combination of many decision trees, which is mainly used for classification, regression, and possible other tasks [35, 42, 50]. Several decision trees are combined as so-called ”weak learners” in a random forest to train datasets [35]. For example, in classification, each individual tree in the forest decides on its own. Then the votes of the trees are counted and the class with the most votes wins the vote [42, 50]. This ML method is characterized by the fact that it trains datasets relatively quickly and simply, since the trees make their decisions independently of each other, thus enabling parallel processing [42]. To use this ML algorithm for image enhancement, the datasets must first be pre-processed [8, 35]. A frequently used method is the Z-Score normalization, which transforms and normalizes the data according to the Ratio rule [35]. However, Z-Score Normalization does not change the size ratio of the individual feature sets to each other and thus does not influence the random forest [35]. Instead, HE methods such as Adaptive Gamma Correction with Weighting Distribution (AGCWD) or Contrast Accumulated Histogram Equalization (CACHE) are applied to preprocess the datasets to improve contrast [35]. In contrast to Z-Score Normalization, HE methods have proven to work well and therefore better [35]. Again, it becomes clear that HE is a classical method that forms the basis for the ML Model.

3.2.4. Support Vector Machine (SVM). The Support Vector Machine (SVM) is known as a supervised ML method for binary classification and regression and is an attractive approach for data modeling [35, 38]. The idea here is to separate the instances of two classes by a linear function in the form of a so-called hyperplane [38]. The optimal hyperplane is the one with the maximum margin, i.e. the largest distance between the two classes [38]. This can be represented in a function using support vectors [38]. In addition to linear classification, SVM can also classify non-linear based on kernel functions [38, 43]. For the preprocessing of datasets for this ML method, Z-Score normalization is conventionally used, as in Random Forest [35]. In addition, HE methods such as AGCWD or CACHE are often used [35]. Furthermore, in medical images, such as computed tomography images, Contrast Limited Adaptive HE (CLAHE) is generally used to improve the contrast and thus process the images in advance [43]. In this case, SVM is used to train various sample features [43]. In contrast to Random Forest, both the conventional Z-Score normalization and HE methods work well for SVM to pre-process the images [35].

3.3. Application areas

ML based computer vision approaches to image enhancement are becoming more and more relevant and important [16, 17, 24, 51]. One reason for this is the broad spectrum of possible applications of the individual forms of development. In the following, the areas of application recognizable from the research are presented and the relevance is illustrated by means of examples.

3.3.1. Low-Light Image Enhancement. The areas of application of image enhancement are as diverse as the methods already presented. The most frequently mentioned application area is that of low-light image enhancement [14, 18, 20, 22]. In this case, little light during acquisition leads to image noise and low contrast, which makes further processing with computer vision methods such as image splitting, object recognition or tracking difficult [18, 20, 22, 22]. This is where image enhancement comes in and aims to make the images both more visually appealing and more suitable for the mentioned computer vision [18, 22]. This becomes clear, for example, in the domain of monitoring public places, traffic management and driver assistance systems [14]. It is not appropriate to simply brighten the low-light image, since the brighter domains will be saturated, which in turn leads to a loss of detail [22]. In this field of application, HE and Retinex are used, which are increasingly applied in addition to CNN-based methods [14].

3.3.2. Rainy Image Enhancement. In addition to insufficient light during image acquisition, rain can also severely impair the visibility of outdoor systems [26, 52]. Due to the increasing interest in improving images with rain impairments, numerous methods have been proposed in recent years [26, 52]. Some work concentrates on eliminating rain in video sequences, while others try to remove the effects of rain on a single image [26, 53]. In both cases, the main goal is to remove rain streaks from the images [26, 53]. Since the previous methods have had varying degrees of success, [26] presented a solution that uses the so-called Dark Channel Prior method to remove the rain stripes first and then the fog [26].

3.3.3. Image Dehazing. Images shot in a hazy environment such as haze, fog, or smoke often suffer a significant loss of quality in the form of low contrast or color change [21, 24, 52, 53]. This occurs because photons are reflected, scattered, and absorbed by the object surface as they move in the direction of the camera [21, 24]. As a result, it is more difficult for the human eye or outdoor computer vision systems to detect individual features of the objects in the images [24, 53]. For this reason, Image Dehazing is an important topic in the computer vision domain and has been extensively researched in recent years [24]. In contrast to the classical image enhancement methods of noise reduction and contrast enhancement, Image Dehazing aims to reduce or even remove the haze in the image to achieve satisfactory visual results [24]. This is achieved by various methods of HE, such as Adaptive HE or Contrast Limited Adaptive HE or by Retinex-based algorithms [24].

3.3.4. Lane Detection. According to the European Accident Research and Safety Report 2013, more than 90% of driving accidents are caused by human error [37]. As a result, there is great potential for the application of ML based computer vision approaches to image enhancement in connection with lane detection on roads [37]. An example is the ”fast learning method” based on CNNs using the Extreme Learning Machine (ELM) by [37]. In general, lane recognition consists of preprocessing, detection, and tracking. The non-improved input images are often considered unreliable due to the complexity and illumination in road scenes, for example in cases with shadows, occlusions, and curves [37]. CNNs are implemented to enhance the input images and extract the regions of interest (ROIs) before application to ensure optimal input to the lane assistance software [37].

3.3.5. Underwater Research. In recent years, underwater research has become more active [54]. For example, through the study of marine species, the exploration of wrecks, the inspection of underwater cables and pipelines, the analysis of underwater scenes and the search and rescue of human lives

[17, 55]. Due to the often-limited time and the limited functionality of the underwater technology, image recordings are used to carry out the activities. Scattering and absorption occurs under water and the images normally have a weak clarity, weak contrast, blurred or cloudy details and color variations [17, 55]. [17] addressed these problems using the Multi-Scale Retinex Framework with advanced filters and offer approaches for the use of ML.

3.3.6. Military. Image enhancement is also a crucial technique in the domain of infrared filters [6]. Infrared images are mainly used in the military domain and are of enormous relevance for area surveillance and target acquisition [6]. Low contrast, noise, blurred image boundaries or compression artifacts are unavoidable in images captured with infrared technology [6]. Therefore, [6] have developed an image enhancement method based on HE and supplemented by a so-called ”cellular automaton”. This combined application turned out to be a very effective method for improving infrared images [6].

3.3.7. Manuscript Improvement. Ancient, discovered manuscripts are essential for mankind to reconstruct human history and to understand what and especially why events happened in the past. Since the records of the past are usually several thousand years old, they show considerable damage and are therefore hardly legible and must be improved for examination [41]. The damages include, for example, deterioration of contrast and general quality of the paper, which contribute to the unrecognizable nature of the characters [41]. Therefore, [41] propose a method for image enhancement based on the Ratio Rule algorithms in combination with the ”K Nearest Neighbors” (KNN) algorithms. This application restores the natural colors of the manuscript [41]. The method described does not lead to an improvement in the quality of the manuscript, since the paper logically remains damaged, but it does lead to a significant improvement in the readability [41].

3.3.8. Medical diagnosis. Liver cell carcinoma is the most common form of liver cancer caused by cirrhosis of the liver [35]. Approximately 60% of liver tumors are diagnosed at a later stage with a survival time of 6 months [35]. In contrast, tumors detected at an early stage have a 5-year survival rate of 70% [35]. For this reason, image enhancement plays a very important role in the medical diagnosis domain [8, 35, 43, 56–58]. The most used diagnostic method for hepatocellular carcinoma is computed tomography (CT), which cannot be used to confirm the diagnosis alone due to gray scale intensity factors [35]. To counteract this problem, CT images are enhanced by HE methods such as Contrast Limited Adaptive HE or Modified Histogram-Based Contrast Enhancement. These HE methods are becoming more and more important nowadays and are therefore frequently used to improve medical images [8, 35, 58]. After the images have been successfully enhanced, they can be trained using the Support Vector Machine or decision trees to classify the different sample features [7, 8, 35, 57]. In addition to the classical algorithms, DL methods such as CNN are used to classify the features [8]. These ML methods are currently popular for the analysis of medical datasets since the classifiers already learned can also be used to diagnose new patients [7]. For this reason, ML has become indispensable in the domain of medical diagnostics [7].

3.3.9. 3D Reconstruction. 3D reconstruction is an increasingly important topic in computer vision with numerous applications, for example in the domain of robotics or augmented reality (AR) [51]. Images captured by an RGB-D camera often have too low a resolution for 3D reconstruction because they are noisy and incomplete, which is where image enhancement comes in to improve the quality of the camera image [51]. Numerous researchers have investigated how DL-based methods can be used to improve the resolution of depth images [51, 59]. For example, Riegler et al. [59] proposed a method that combines the advantages of DL with a variable method to restore the high-resolution depth image [51]. Consequently, [51] propose a new method implemented in the DL Framework to achieve a satisfactory depth improvement.

4. Discussion

In the field of ML-based computer vision approaches to image enhancement, future attention will be focused on increasing efficiency in terms of enhancing individual images, increasing flexibility of algorithms, as well as decreasing power requirements on hardware. When image enhancement is to be performed, it cannot be directly assumed that the application used will also produce the desired results [16, 26, 28, 41]. This is due to the heterogeneous initial situation of the individual use cases. Furthermore, it is often unclear in advance which methodology will provide the best result, and time-consuming experimentation with different methods is used to find the optimal solution [16, 26, 28, 41]. Potential time savings can be seen in terms of increasing efficiency in selecting the appropriate methods. A large number of the presented methods work excellently in their specific application area, but they are only designed for this one purpose and usually cannot be applied to different problems. Thus, there are a large number of methods that solve specific problems well, but consequently are not adaptable and transferable to other application domains. This can be proved, for example, with an algorithm which was developed for color images and therefore has a different basic framework than an algorithm for infrared images [6, 16, 28, 49, 51]. Therefore, there is a great potential to develop algorithms in the future that cover a wide range of applications, are adaptable, and can independently solve different problems.

The methods presented in this literature review in the area of ML-based computer vision approaches for image enhancement often have the problem that they place high demands on hardware resources and are thus not usable for the masses [28, 37, 49]. This is especially the case when real-time applications are involved [37]. Among other reasons, this is due to the fact that the individual algorithms have to perform elaborate computations in order to be able to deliver the desired results. These high system requirements can severely limit the application, but this can be compensated for by advances in computer systems, although it will also be a task in the future to make applications more resource-efficient without losing speed and effectiveness. The issues described should be adequately addressed in the future to ensure progress in a forward-looking area of computer vision. As the field is developing rapidly, such as in autonomous driving, the development of the addressed areas turns out to be a basic building block for a technological future [60].

4.1. Limitations

The limitation of this study was often the transparency of the works to be compared, as not all of them provided the complete information needed. Moreover, it was limited to the selected essential and, from our point of view, most relevant procedures and models, which means that in the future it may be necessary to analyze other and innovative procedures, which is very likely due to the immense development of the research field, also related to the progress of technology in the field of image recording.

5. Conclusion

We considered all major peer-reviewed journals and conference papers - supplemented by a forward and backward search - on ML-based computer vision approaches to image enhancement and conducted a comprehensive literature review. First, the basic, classical image enhancement methods were considered, then the relevant ML models, and finally the potential applications of these. The structure results from the fact that in practice the ML methods are often based on the classical methods, evident among others in the example of HE and Support Vector Machine and Random Forest, respectively. It also highlighted research gaps for future research in this area. Since we limited this analysis to peer-reviewed sources, it is possible that potentially interesting work was not considered in this research. Future work should systematically address the research gaps formulated.

References

[1] S. S. Agaian, K. Panetta, and A. M. Grigoryan, “A new measure of image enhancement,” in *IASTED International Conference on Signal Processing & Communication*. Citeseer, 2000, pp. 1–7.

[2] S. Rajput and S. Suralkar, “Comparative study of image enhancement techniques,” *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 1, pp. 11–21, 2013.

[3] M.-J. Seow and V. K. Asari, “Ratio rule and homomorphic filter for enhancement of digital colour image,” *Neurocomputing*, vol. 69, no. 7-9, pp. 954–958, 2006.

[4] A. Mehra, M. Mandal, P. Narang, and V. Chamola, “ReViewNet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions,” vol. 22, no. 7, pp. 4256–4266, 2021.

[5] V. Voronin, A. Zelensky, and S. Agaian, “3-d block-rooting scheme with application to medical image enhancement,” vol. 9, pp. 3880–3893, 2021.

[6] W. Qi, J. Han, Y. Zhang, and L. fa Bai, “Infrared image enhancement using cellular automata,” *Infrared Physics & Technology*, vol. 76, pp. 684–690, 2016.

[7] I. Kononenko, “Machine learning for medical diagnosis: history, state of the art and perspective,” *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89–109, 2001.

[8] J. Latif, C. Xiao, A. Imran, and S. Tu, “Medical imaging using machine learning and deep learning algorithms: A review,” in *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 2019.

[9] I. E. Naqa and M. J. Murphy, “What is machine

learning?" in *Machine Learning in Radiation Oncology*. Springer International Publishing, 2015, pp. 3–11.
[10] A. Lindner and S. Susstrunk, "Semantic-improved color imaging applications: It is all about context," *IEEE Transactions on Multimedia*, pp. 1–1, 2015.
[11] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.
[12] G. Abramo and C. A. D'Angelo, "Evaluating research: from informed peer review to bibliometrics," *Scientometrics*, vol. 87, no. 3, pp. 499–514, 2011.
[13] D. W. Aksnes, L. Langfeldt, and P. Wouters, "Citations, citation indicators, and research quality: An overview of basic concepts and theories," *SAGE Open*, vol. 9, no. 1, pp. 1–17, 2019.
[14] Y. Guo, X. Ke, J. Ma, and J. Zhang, "A pipeline neural network for low-light image enhancement," *IEEE Access*, vol. 7, pp. 13 737–13 744, 2019.
[15] J. Kim, S. Song, and S.-C. Yu, "Denoising auto-encoder based image enhancement for high resolution sonar image," in *2017 IEEE Underwater Technology (UT)*, 2017.
[16] W. Qi, J. Han, Y. Zhang, and L. fa Bai, "Hierarchical image enhancement," *Infrared Physics & Technology*, vol. 76, pp. 704–709, 2016.
[17] S. Zhang, T. Wang, J. Dong, and H. Yu, "Underwater image enhancement via extended multi-scale retinex," *Neurocomputing*, vol. 245, pp. 1–9, 2017.
[18] Y. Ren, Z. Ying, T. H. Li, and G. Li, "LECARM: Low-light image enhancement using the camera response model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 4, pp. 968–981, 2019.
[19] X. Wang, Z. An, H. Wang, and Y. Chang, "MR brain image enhancement via learning ensemble," in *2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, 2018.
[20] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1032–1041, 2013.
[21] Y. Cho, J. Jeong, and A. Kim, "Model assisted multi-band fusion for single image enhancement and applications to robot vision," *IEEE Robotics and Automation Letters*, pp. 1–1, 2018.
[22] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 982–993, 2017.
[23] J. Ma, X. Fan, J. Ni, X. Zhu, and C. Xiong, "Multi-scale retinex with color restoration image enhancement based on gaussian filtering and guided filtering," *International Journal of Modern Physics B*, vol. 31, no. 16-19, p. 1744077, 2017.
[24] W. Wang and X. Yuan, "Recent advances in image dehazing." *IEEE CAA J. Autom. Sinica*, vol. 4, no. 3, pp. 410–436, 2017.
[25] T. Celik and T. Tjahjadi, "Automatic image equalization and contrast enhancement using gaussian mixture modeling," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 145–156, 2012.
[26] Z. Shi, Y. Li, M. Zhao, Y. Feng, and L. He, "Multi-stage filtering for single rainy image enhancement," *IET Image Processing*, vol. 12, no. 10, pp. 1866–1872, 2018.
[27] A. Temizel, "Image resolution enhancement using wavelet domain hidden markov tree and coefficient sign estimation," in *2007 IEEE International Conference on Image Processing*, 2007.
[28] M. Z. Zhang, M.-J. Seow, L. Tao, and V. K. Asari, "A tunable high-performance architecture for enhancement of stream video captured under non-uniform lighting conditions," *Microprocessors and Microsystems*, vol. 32, no. 7, pp. 386–393, 2008.
[29] J. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 889–896, 2000.
[30] M. Abdullah-Al-Wadud, M. Kabir, M. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, 2007.
[31] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1143–1152, 2006.
[32] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.
[33] T. K. Kim, J. K. Paik, and B. S. Kang, "Contrast enhancement system using spatially adaptive histogram equalization with temporal fil-

tering," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 1, pp. 82–87, 1998.

[34] S.-C. Huang and C.-H. Yeh, "Image contrast enhancement for preserving mean brightness without losing image features," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 5-6, pp. 1487–1492, 2013.

[35] A. Sada, Y. Kinoshita, S. Shiota, and H. Kiya, "Histogram-based image pre-processing for machine learning," in *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*, 2018.

[36] U. K. Nafis, K. Arya, and M. Pattanaik, "Histogram statistics based variance controlled adaptive threshold in anisotropic diffusion for low contrast image enhancement," *Signal Processing*, vol. 93, no. 6, pp. 1684–1693, 2013.

[37] J. Kim, J. Kim, G.-J. Jang, and M. Lee, "Fast learning method for convolutional neural networks using extreme learning machine and its application to lane detection," *Neural Networks*, vol. 87, pp. 109–121, 2017.

[38] M. Hearst, S. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[39] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, no. 4, pp. 525–533, 1993.

[40] T. Remez, O. Litany, R. Giryes, and A. M. Bronstein, "Deep class-aware image denoising," in *2017 International Conference on Sampling Theory and Applications (SampTA)*, 2017.

[41] R. Kataria, T. Sharma, T. Choudhury, and P. Kumar, "Image enhancement of old manuscripts using machine learning," in *2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, 2017.

[42] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[43] L. Ali, K. Khelil, S. K. Wajid, Z. U. Hussain, M. A. Shah, A. Howard, A. Adeel, A. A. Shah, U. Sudhakar, N. Howard, and A. Hussain, "Machine learning based computer-aided diagnosis of liver tumours," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing*, 2017, pp. 1–7.

[44] M. Klaiber, D. Sauter, H. Baumgartl, and R. Buettner, "A systematic literature review on transfer learning for 3d-CNNs," in *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021.

[45] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *Twenty-second international joint conference on artificial intelligence*, 2011, pp. 1–6.

[46] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 1733–1740.

[47] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611–629, 2018.

[48] M. Nandhini, A. Anand, and R. Kumar, "Analysis on domain based method for image enhancement," *International Journal of Application or Innovation in Engineering & Management (IJAIEM), ISSN*, pp. 2319–4847, 2012.

[49] M. Z. Zhang, M.-J. Seow, and V. K. Asari, "A high performance architecture for color image enhancement using a machine learning approach," *International Journal of Computational Intelligence Research-Special Issue on Advances in Neural Networks*, vol. 2, no. 1, pp. 40–47, 2006.

[50] R. Buettner, S. Sauer, C. Maier, and A. Eckhardt, "Towards ex ante prediction of user performance: A novel NeuroIS methodology based on real-time measurement of mental effort," in *2015 48th Hawaii International Conference on System Sciences*, 2015.

[51] J. Li, W. Gao, and Y. Wu, "High-quality 3d reconstruction with depth super-resolution and completion," *IEEE Access*, vol. 7, pp. 19 370–19 381, 2019.

[52] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[53] L.-W. Kang, C.-W. Lin, and Y.-H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1742–1755, 2012.

[54] P. Zhuang, C. Li, and J. Wu, "Bayesian retinex underwater image enhancement," *Engineering Applications of Artificial Intelligence*, vol. 101, p. 104171, 2021.

[55] P. Sahu, N. Gupta, and N. Sharma, "A survey on underwater image enhancement techniques," *International Journal of Computer Applications*, vol. 87, no. 13, pp. 1–5, 2014.

[56] R. Gupta and S. Sehgal, “Gray-scale image enhancement using differential evolution optimization algorithm,” in *2016 International Conference on Information Technology (InCITe) - The Next Generation IT Summit on the Theme - Internet of Things: Connect your Worlds*, 2016.

[57] L. Nanni, A. Lumini, and S. Brahnam, “Local binary patterns variants as texture descriptors for medical image analysis,” *Artificial Intelligence in Medicine*, vol. 49, no. 2, pp. 117–125, 2010.

[58] Suharjito, N. Thiracitta, H. Gunawan, and G. Witjaksono, “The comparison of some hidden markov models for sign language recognition,” in *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)*, 2018.

[59] G. Riegler, M. Rüther, and H. Bischof, “ATGV-net: Accurate depth super-resolution,” in *Computer Vision – ECCV 2016*. Springer International Publishing, 2016, pp. 268–284.

[60] B. H. Guo, Y. Zou, Y. Fang, Y. M. Goh, and P. X. Zou, “Computer vision technologies for safety science and management in construction: A critical review and future research directions,” *Safety Science*, vol. 135, p. 105130, 2021.