

Received January 29, 2022, accepted February 25, 2022, date of publication March 2, 2022, date of current version March 22, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3156273

# Single-Image Reflection Removal Using Deep Learning: A Systematic Review

ALI AMANLOU<sup>1</sup>, AMIR ABOLFAZL SURATGAR<sup>1</sup>, (Senior Member, IEEE), JAFAR TAVOOSI<sup>2</sup>,  
ARDASHIR MOHAMMADZADEH<sup>3</sup>, AND AMIR MOSAVI<sup>4,5,6,7</sup>, (Student Member, IEEE)

<sup>1</sup>Distributed and Intelligent Optimization Research Laboratory, Electrical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran 15916 34311, Iran

<sup>2</sup>Department of Electrical Engineering, Ilam University, Ilam 67187 73654, Iran

<sup>3</sup>Department of Electrical Engineering, University of Bonab, Bonab 55517 61167, Iran

<sup>4</sup>Faculty of Civil Engineering, TU-Dresden, 01069 Dresden, Germany

<sup>5</sup>John von Neumann Faculty of Informatics, Obuda University, 1034 Budapest, Hungary

<sup>6</sup>Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology in Bratislava, 811 07 Bratislava, Slovakia

<sup>7</sup>Institute of Information Society, University of Public Service, 1083 Budapest, Hungary

Corresponding authors: Amir Abolfazl Suratgar (a-suratgar@aut.ac.ir) and Amir Mosavi (amir.mosavi@mailbox.tu-dresden.de)

**ABSTRACT** Images captured through the glass often consist of undesirable specular reflections. These reflections detected in front of the glass remarkably reduce the quality and visibility of the scenes behind it. The process of reflection removal from images through the glass has many important applications in computer vision projects. Recently deep learning-based methods are being utilized for reflection removal so widely. In this article, we proposed a systematic literature review on the topic of single-image reflection removal using deep learning methods which were published between the years 2015 to 2021. A total number of 1600 research papers were extracted from five different online databases and digital libraries (IEEE Xplore, Google Scholar, Science Direct, SpringerLink and ACM Digital Library). After following the study selection procedure, 25 research papers were selected for this systematic review. The selected research papers were then analyzed to answer 7 key research questions that we have come up with to comprehensively explore the use of deep learning and neural networks for single-image reflection removal. After reading this article, future researchers will have a solid idea in the research field and will be able to work on their own research. The results provided in this proposed systematic review illustrate the main challenges that are encountered by researchers in this field and recommend encouraging directions for future research work. This review will also be helpful for researchers in discovering accessible datasets that can be used as benchmarks for comparing their proposed deep learning techniques with other studies in this research area.

**INDEX TERMS** Deep learning, reflection removal, reflection separation, systematic review.

## I. INTRODUCTION

When we capture images through transparent material especially glass, the images frequently consist of some undesirable reflections. These reflections will lower the quality and visibility of the images [1], [2]. Some photographers may make a dark condition or modify the location of the camera but this approach is not efficient for reflection removal due to the limitations on space. These reflections have two main problems: first, they will lower the quality of the scenes behind the glass and second, they have some effects on the results of applications like segmentation or classification [1]. Therefore, this problem may cause most computer vision

algorithms presumably to fail. The aim of reflection removal is to improve the visibility of the scene behind the glass while removing the reflections. Reflection removal is one of the most challenging projects in computer vision and this originates from its ill-posed character. The features of the background and reflections are so much similar. Therefore, these similarities will make it more complicated to remove the reflections and retrieve the scene behind the glass [2]. Although it is easy for the human vision system (HVS) to identify and distinguish the transmitted scene from reflection but it is a complicated task for computers even if blurry and ghosting artifacts are easy to notice [5]. In recent years, many approaches are proposed as a solution for reflection removal and as it is shown in figure 1, they are categorized into two groups: conventional mathematical methods

The associate editor coordinating the review of this manuscript and approving it for publication was Tomasz Trzcinski<sup>1</sup>.

(non-learning methods) and deep learning-based algorithms. Non-learning methods used handcrafted priors as a solution for this problem. [3]. Many of these proposed non-learning-based reflection removal algorithms work under special circumstances. In other words, these non-learning solutions are efficient when their assumptions are fully satisfied [2]. Most of the time these non-learning methods are based on the presumption that the image taken through the glass is a linear combination of transmitted scene and reflection [5].

In some conditions, when the design, patterns and properties of the background scene show similarities to those of the reflections, it is difficult for non-learning-based algorithms to remove the reflections and retrieve the background at the same time [4]. And also when a non-learning algorithm tries to solve the linear equation of background **B** and reflection **R** without any prior knowledge, it will face an infinite number of solutions for the processing of the single image [5]. Reflection removal approaches could also fall into 2 main categories: Multiple image reflection removal and single-image reflection removal. Multiple image reflection removal approaches can be classified into four different groups:

- Multiple polarized images
- Focus and defocus image pairs
- Flash and non-flash image pairs
- Video sequences

More details are shown in figure 2. Mobile phone users and photographers tend to take a single image instead of taking a series of images because multiple image reflection removal may not be possible due to the limitations. Many different models have been proposed to remove the reflections and these models are efficient to some extent but they still have limitations [5].

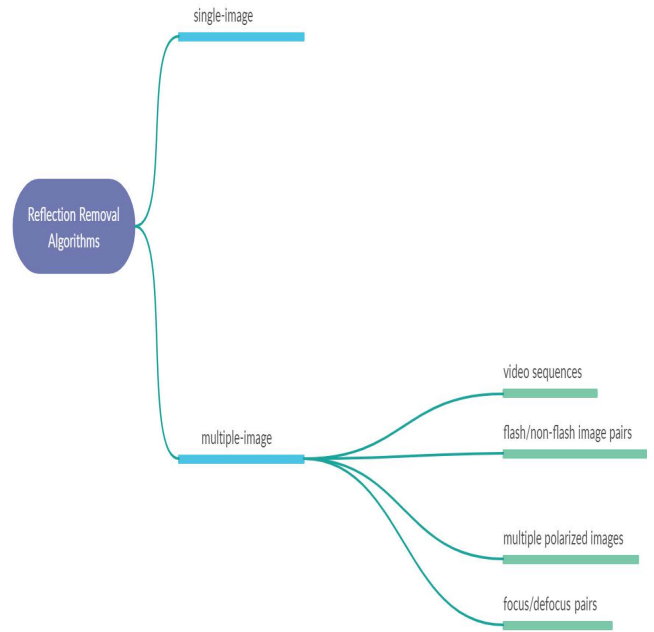


FIGURE 2. Classification of reflection removal algorithms.

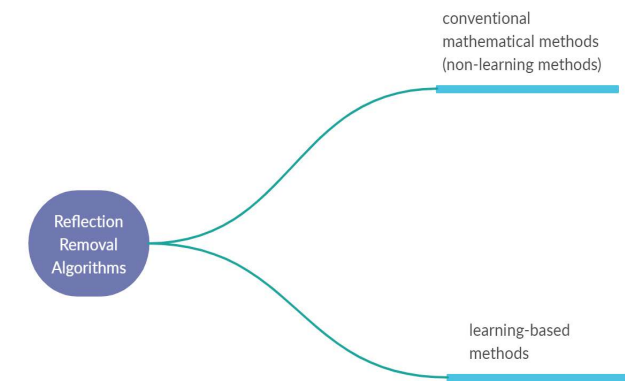


FIGURE 1. Classification of reflection removal algorithms (learning and non-learning methods).

According to the promising results of deep learning in high-level and low-level computer vision tasks, its detailed modeling capability also profits reflection removal projects. Recently deep learning-based algorithms are being proposed to capture the reflection properties more efficiently and comprehensively. Deep learning-based algorithms which have been proposed recently, illustrate enhanced modeling

capability that detects a wide range of reflection image characteristics [4]. Deep learning algorithms rely on mapping images to high-dimensional features using deep learning techniques. On other hand, conventional methods work based on handcrafted features and complex mathematical analysis like edges and gradient [5].

Fan *et al.* [6] presented the first approach which uses a convolutional neural network (CNN) for the single-image reflection removal problem. With two cascaded networks, they first utilized a CNN to predict edges of the background layer. This predicted edge is used as guidance for reconstructing the background layer by the second CNN. The semantic structure is not considered because they use pixel-wise loss function in the procedure of training. Generative Adversarial Networks (GAN)-based algorithms and techniques have also shown distinguished results in reflection removal problems. PL Net [7] was proposed by Zhang *et al.* which is trained by exclusion loss, adversarial loss and loss function composed of feature loss. Their network and loss function concentrate on both low and high-level image information. Yang *et al.* [8] proposed a method that is based on the prediction of the reflection layer and background layer. ERR Net [9] was proposed by Wei *et al.* which is able to be trained by misaligned data. [1]. In addition to these papers that we mentioned in the current paragraph, many other research papers are explored in this systematic review to find comprehensive answers for our proposed questions.

In the end, our proposed Systematic Literature Review (SLR) intends to collect, analyze and introduce research papers in the field of single-image reflection removal using deep learning. Our systematic review also illustrates research directions for new researchers in the reflection removal

field by emphasizing weak areas of existing reflection removal algorithms based on deep learning that needs further explorations. In this SLR, we also examined available datasets/databases used by researchers for validation of their proposed algorithms. Different algorithms of deep learning for single-image reflection removal were also examined. Due to the fact that the number of published research papers in the area of reflection removal task is huge, we limit this SLR to address only research papers that utilized deep learning algorithms. Therefore, our paper is the first systematic literature review that focuses on the topic of single-image reflection removal using deep learning techniques.

The rest of this systematic review is arranged as follows. Section II is a discussion of the utilized research methodology in this SLR. This section includes planning, conducting and reporting. Section III is a comprehensive analysis of the considered studies which discusses the results and findings of our studies. Our SLR is concluded in section IV and section V is acknowledgment. Finally, section VI is the references.

## II. RESEARCH METHODOLOGY

The review conducted in this article is based on a systematic review presented by Kitchenham and Charters method [10]. A Systematic Literature Review (SLR) is a very comprehensive method for evaluation and interpretation of all available research related to a particular research question or field. SLR is a well-defined, trustworthy, rigorous, and auditable way of presenting a fair assessment of a research topic [10]. In this systematic literature review, we followed Kitchenham and Charters [10] guidelines. As the presented method in Kitchenham and Charters recommends, this SLR has been done in three phases: Planning, conducting and reporting. Each phase is divided into several sub-sections that we have explained below.

### A. PLANNING

#### 1) A NEED FOR SYSTEMATIC REVIEW

Early literature reviews show that to the best of our information, no systematic literature review had been published so far in the field of single-image reflection removal using deep learning. Furthermore, our conducted survey is one of the earliest reviews conducted so far in this field. Therefore, we believe it is the best time to collect and analyze the available related research projects in the domain.

#### 2) DEVELOPMENT OF THE RESEARCH PROTOCOL

The review protocol is the most significant difference between a systematic review and a traditional review [10]. In the process of conducting this systematic review, we searched for related articles in several databases. Secondly, we reduced the number of articles by applying quality assessments, inclusion and exclusion criteria. Finally, some research questions were used to propose this SLR.

#### 3) RESEACH QUESTIONS

The major objective of our proposed systematic review is to identify and inspect research papers that used deep learning

for single-image reflection removal. Related research papers were searched in various databases to reach this objective and the following research questions arose. Table 1 presents the research questions and their descriptions.

TABLE 1. Research questions.

No.	Questions	Descriptions
RQ1	What is the distribution of the selected research articles over the last seven years and their types?	Answering this question will give an overview of the type and publication year of considered papers in Reflection Removal.
RQ2	What datasets/databases are used to train the proposed deep learning model in considered publications?	Available datasets/databases with enough data are one of the most considerable factors in training deep learning algorithms.
RQ3	What datasets/databases are used for conducting experiments on single-image reflection removal?	Available datasets/databases with enough data are one of the most considerable factors in testing deep learning algorithms.
RQ4	What are the architectures of the proposed reflection removal networks in each paper?	Answering this question will give researchers an overview of the latest trends in applying deep learning algorithms in reflection removal.
RQ5	What quantitative metrics have been used to accredit the efficiency of the proposed deep learning techniques?	Answering this question will give researchers an overview of the utilized quantitative metrics in the field.
RQ6	What are the current limitations and challenges in this research field?	Answering this question will guide new researchers in potential future projects.
RQ7	What are the possible future works and directions in the single-image reflection removal field via deep learning techniques?	Answering this question will guide new researchers in potential future projects.

#### 4) INCLUSION AND EXCLUSION CRITERIA

Raising some inclusion and exclusion criteria guarantees that merely research papers that are related to our topic are being studied. We set up 5 inclusion and 5 exclusion criteria. The inclusion, exclusion are shown in the following Tables 2 and 3. Our search based on keywords end in 1600 research papers relevant to our topic. Note that we removed all of the duplicate papers at first. After a complete review of the research papers, we included papers according

to the criteria mentioned in Table 2. Then we excluded the rest of the papers according to the criteria mentioned in Table 3. Eventually, we removed some papers based on quality assessment rules which are explained later in this paper.

**TABLE 2. Inclusion criteria.**

Criteria	Description
1	Papers that are published between 2015 – 2021.
2	Papers that are published in conferences and journals
3	Papers that are using deep learning and neural networks as the main method for reflection removal.
4	Papers that are proposing networks for single-image reflection removal.
5	Peer-reviewed research papers

**TABLE 3. Exclusion criteria.**

Criteria	Descriptions
1	Papers that have unrelated titles to our research topic.
2	Papers that do not present clear information and methodology.
3	Papers that are not published in the English Language.
4	Papers that are not related to our research field.
5	Papers that are not using deep learning and neural networks for reflection removal problem

## 5) QUALITY ASSESSMENT RULES

Quality Assessment Rules (QARs) is a concept based on the principle to decide about the overall quality of the selected research articles [59]. Applying these criteria will be helpful to select the most relevant research articles with high quality. The 10 quality assessment rules applied to the selected research articles are mentioned in Table 4. Each of these QARs is worth 1 point out of 10. We determined the points of each question as follows: when completely and comprehensively answered points = 1, when the answer is above normal points = 0.75, when the answer is average points = 0.5, when the answer is below average points = 0.25, and when the question is not answered at all points = 0. Finally, we add all of these points and the summation shows the points of each research paper. The points of 7 or more express that we included that research paper in this systematic review.

## B. CONDUCTING

In the second phase, we searched online databases and digital libraries using search strings and keywords, in order

**TABLE 4. Quality assessment rules (QARs).**

Criteria	Descriptions
1	Is the research paper appropriately prepared?
2	Are the aims and purposes of the research paper clear? Is the information presented in the research paper related to the aims and purposes of our systematic literature review?
3	Is the information presented in the research paper well-explained and clear?
4	Are the utilized deep learning models explained completely?
5	Are different sections of the research paper related to each other?
6	Is the utilized dataset/database identified completely?
7	Are any conducted experiments included in the research paper?
8	Is the conducted experiment justifiable?
9	Does the paper present a comparison of results with previous research papers?
10	Is it an effective and useful research paper in the single-image reflection removal field?

to find the most related articles and research studies for systematic review. This phase includes extracting research articles and selecting those according to inclusion and exclusion criteria and also quality assessment rules. More details are illustrated in figure 3.

The first step in conducting phase is to search through digital libraries, databases and scientific search engines in order to explore the most relevant articles. For this aim, we searched through the most popular online databases and digital libraries such as IEEE Xplore, Science Direct, Springerlink, Google scholar and ACM Digital Library. The terms selected for this search were:

- **Deep Learning**
- **Neural Network**
- **Convolutional Neural Networks**
- **Single-image Reflection Removal**

With the keywords defined above, multiple keywords and search strings that are made with logical operators were constructed in order to enhance the search results.

The multiple keywords and search strings used for the SLR are mentioned below:

- “reflection removal” AND “deep learning”
- “reflection removal” AND (“deep learning” OR “neural network”)
- “reflection removal” AND “CNN”

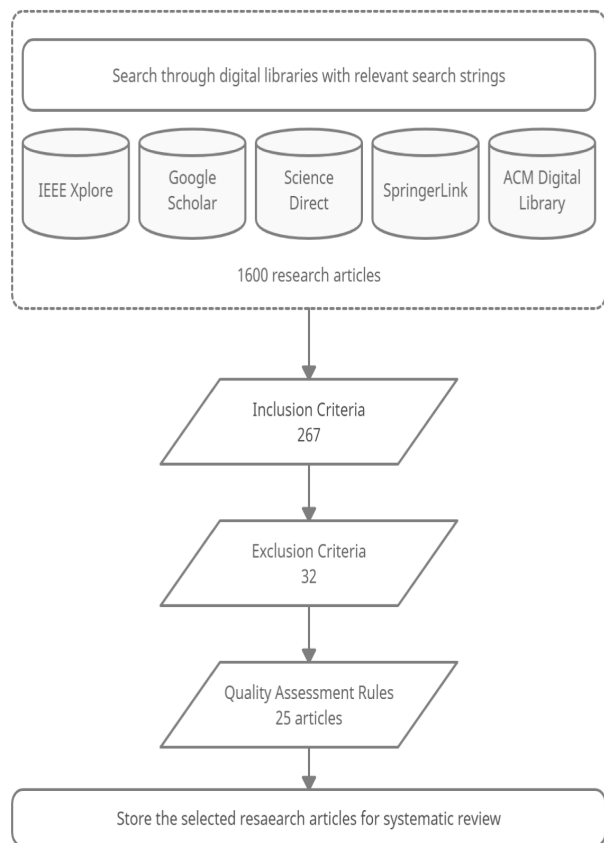


FIGURE 3. The process of searching and selecting research articles.

- “reflection removal” AND “convolutional neural network”
- “reflection removal” AND “deep neural network”
- “reflection separation” AND “deep learning”
- “reflection separation” AND “deep neural network”
- “reflection separation” AND “convolutional neural network”
- “reflection separation” AND “CNN”

Moreover, we manually searched through other sources including reference lists and citations of the selected studies. The publications based on search strings and keywords were made on December 9, 2021, totaling 1600 research papers. As it is illustrated in figure 3, after applying inclusion and exclusion criteria, quality assessment rules and removing duplicate items, finally there were remaining 25 research papers. These 25 papers chosen for this SLR are dependable as they all existed in high-impact digital libraries. The authors of these 25 papers have a significant reputation in the reflection removal research field. On the other hand, these selected papers have a good number of citations. The citation count of selected research papers is shown in figure 4.

C. REPORTING

The last and final stage of our proposed systematic literature review is reporting the results. In this final step, we studied

all the considered research articles comprehensively and analyzed them in order to answer the 7 key research questions that are mentioned in Table 1. Finally, we report our findings in section III which is results and findings. The analyzed and considered research articles and the number of their citation (according to the Google Scholar database) are mentioned in Table 5. The given ID in Table 5 will be utilized whenever one of these papers is needed to be mentioned for the rest of our systematic review.

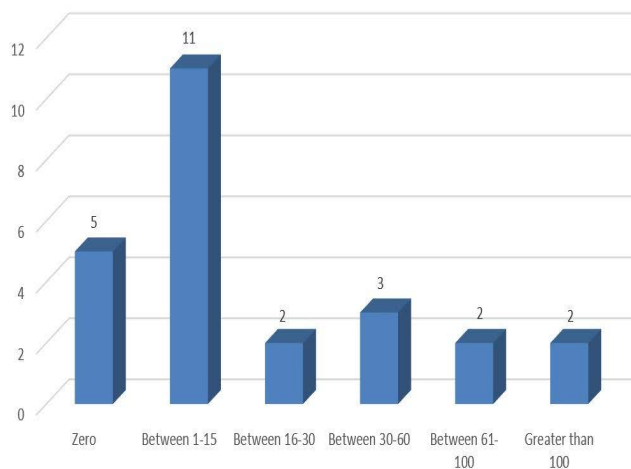


FIGURE 4. Citation count of selected research papers.

III. RESULTS AND FINDINGS

After reviewing and examining all the related research articles, we provide the following answers to all research questions which were mentioned in Table 1.

A. RQ1. WHAT IS THE DISTRIBUTION OF THE SELECTED RESEARCH ARTICLES OVER THE LAST SEVEN YEARS AND THEIR TYPES?

In order to answer question RQ1, we analyzed and examined all the 25 research articles with a high focus on the distribution of publications per year. Figure 5 illustrates the number of published research articles between 2015 and 2021. According to Figure 5, there has been a growth in using deep learning techniques for reflection removal in the last seven years. The highest number of articles were published in the year 2019 over the last 7 years. Table 6 shows the publishing year of each selected article in this review.

In figure 6, the number of published research articles is presented per country. Most of the studies were from China with 11 published articles followed by researchers from Singapore with 4, South Korea with 3 and Taiwan with 2 articles. We also found only one research article followed by researchers from Australia, Japan, India and the United States. This fact shows that researchers from these countries have been focusing more on single-image reflection removal via various deep learning techniques.



TABLE 5. Analyzed articles.

ID	Citations google	Author(s)
P1	170	[6]
P2	0	[13]
P3	25	[2]
P4	61	[4]
P5	2	[11]
P6	1	[12]
P7	0	[15]
P8	2	[16]
P9	14	[17]
P10	3	[29]
P11	87	[8]
P12	6	[1]
P13	44	[19]
P14	55	[9]
P15	33	[21]
P16	18	[5]
P17	1	[81]
P18	1	[23]
P19	0	[24]
P20	1	[25]
P21	14	[26]
P22	154	[7]
P23	1	[27]
P24	0	[79]
P25	0	[80]

Figure 7 indicates the study distribution per type. 28% of the articles were published in the journal while 72% were published in conferences. As it is obviously significant, the percentage of published research articles in conferences is larger than the percentage of published research articles in journals. Most of the considered research articles in this systematic review are published in well-known journals and conferences like CVPR, ICCV and ECCV. According to this merit of the selected articles, we strongly believe that this systematic review can be beneficial in order to find the latest trends for future research studies in the field of reflection removal.

Distribution of publications per year

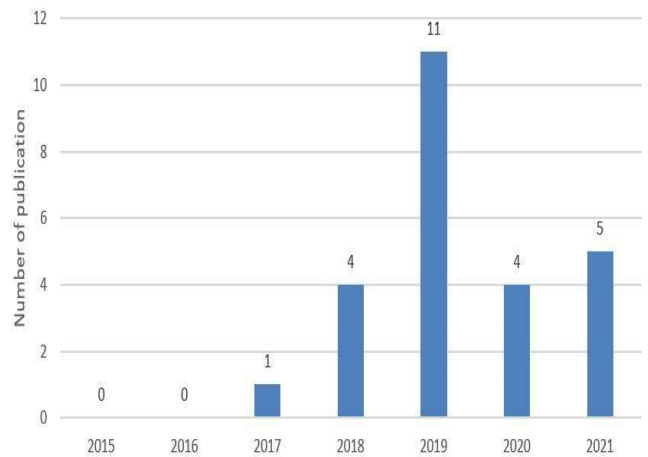


FIGURE 5. Distribution of publications per year.

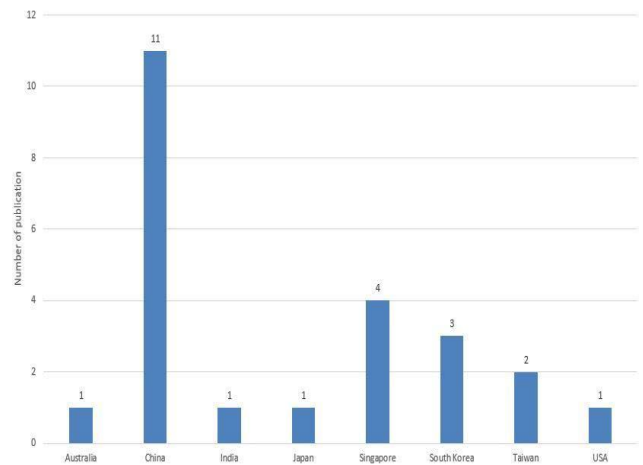


FIGURE 6. Distribution of publications per country.

TABLE 6. Publishing year of each selected articles.

2017	2018	2019	2020	2021
P1	P4, P11, P16, P22	P2, P3, P5, P6, P7, P9, P12, P13, P14, P18, P23	P8, P15, P19, P21	P10, P17, P20, P24, P25

**B. RQ2. WHAT DATASETS/DATABASES ARE USED TO TRAIN THE PROPOSED DEEP LEARNING MODEL IN CONSIDERED PUBLICATIONS?**

To answer the question RQ2, this section provides extracted information from the considered research papers about the datasets used for training the reflection removal algorithms. According to the data-hungry character of deep learning algorithms, the availability of a large dataset containing reflection images is so necessary for the procedure of training the

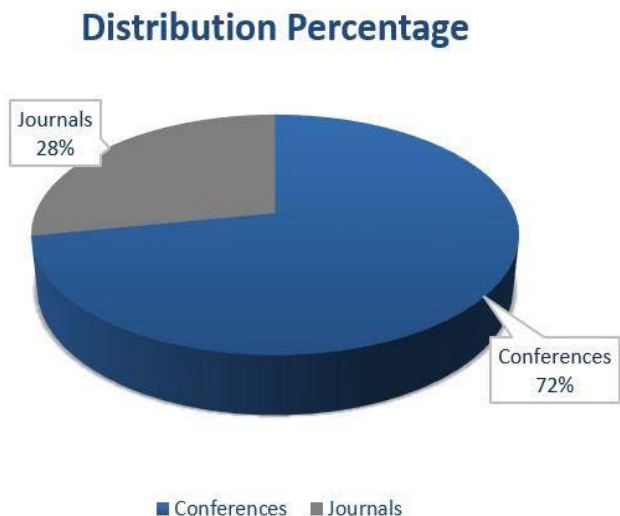


FIGURE 7. Distribution of publications per type.

algorithm [12]. In order to train and test every deep learning technique, many different datasets were used in considered research articles. Most researchers used publicly available datasets to train and test their proposed algorithm. Due to the fact that it is not an easy process to gain a great number of images containing reflections, a substitute way is to synthesize reflection images via two images without reflections as background and reflection [12]. This method has been utilized in a great number of previous mathematical techniques for quantitative evaluations [5], [40], [60], [61], [62]. In the process of synthesizing the dataset, they randomly flipped the images and another image is utilized as a background layer and another image is utilized as the reflection layer [1]. After reviewing all of the considered articles, we summarized the datasets which were used with the purpose of training the deep learning algorithm for each research paper in Table 7. The results of our survey show that researchers tend to use synthetic datasets for training their proposed algorithm rather than using real-world reflection removal datasets. This may be because of the shortage of real-world datasets in this research field. Synthetic datasets have been used in almost all of the selected research papers for training the deep learning technique. In most cases, the synthetic dataset is developed by adding two random images from one or two publicly available datasets. More information is available in Table 7.

**C. RQ3. WHAT DATASETS/DATABASES ARE USED FOR CONDUCTING EXPERIMENTS ON SINGLE-IMAGE REFLECTION REMOVAL?**

To answer the RQ3, we have extracted all of the datasets in which researchers evaluated their proposed deep learning technique and conducted experiments on them in the considered articles. After this process, we have summarized our results in Table 8. In Table 8, the IDs of the publications and the list of benchmarking datasets are mentioned. Due to the extracted information from all of the considered articles,

the experiments mostly use the SIR<sup>2</sup> benchmark dataset to evaluate their proposed deep learning algorithm for single image reflection removal. Single Image Reflection Removal dataset ‘SIR<sup>2</sup>’, proposed by Wan et al. [38] contains a total of  $(20 + 20) \times (7 + 3) \times 3 + 100 \times 3 = 1500$  images. SIR<sup>2</sup> dataset contains 40 controlled indoor scenes with complicated appearance and quality. Each of these 40 scenes consists of a triplet of images (compound image, background and reflection) under three controlled thicknesses of transparent material and seven different depths of field. SIR<sup>2</sup> dataset also contains 100 wild scenes with various thicknesses of transparent material, uncontrolled lightings and different camera settings [38]. These images provided in SIR<sup>2</sup> are all captured via a NIKON D5300 camera with a  $3 \times 10^2$  millimeter lens. The camera settings were all in manual mode. Also, all of the captured images have a resolution of  $1726 \times 1234$ . In the process of capturing a triplet of images in the SIR<sup>2</sup> dataset, Wan et al. [38] captured the mixture image through the transparent material at the first step. The mixture image contains both the background and the reflection. At the second step, they captured just the reflection using a piece of dark paper at the back of the transparent material. At the third step, they captured just the background image by taking the glass away [38]. Due to Table 8 presented for this section, 64% of the considered research articles used the SIR<sup>2</sup> benchmark dataset as one of the main datasets for conducting experiments and evaluating their proposed reflection removal technique.

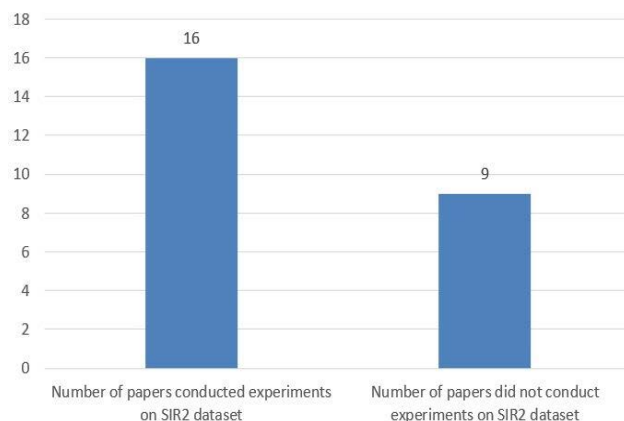


FIGURE 8. Number of papers conducted experiments on SIR<sup>2</sup> dataset.

On the other hand, almost 36% of the considered research articles did not use the SIR<sup>2</sup> dataset at all for doing experiments and evaluation. Almost 16% of the considered research articles only used the SIR<sup>2</sup> dataset for conducting experiments and evaluation. Near 92% of the research articles used real-world datasets, and 8% of the research articles only used synthetic datasets for conducting experiments and evaluation. The results show that most of the researchers tend to use real-world benchmarking datasets for quantitative evaluation. More results are shown in figure 8 and 9.

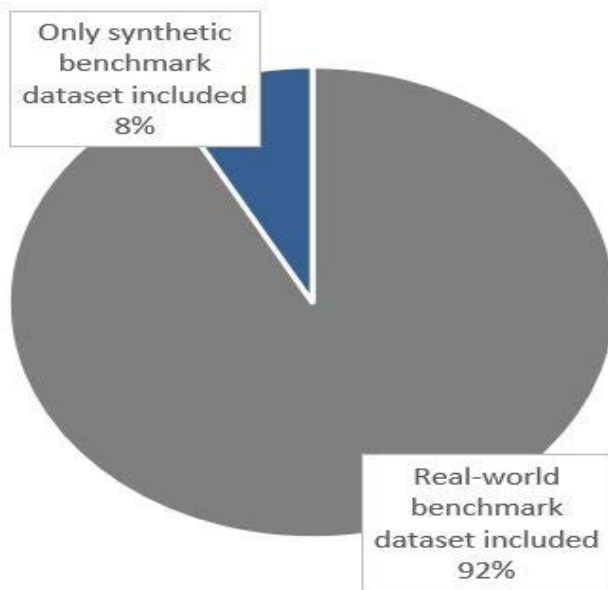
TABLE 7. Utilized training datasets in each selected article.

ID	Author(s)	Training Datasets
P1	[6]	7643 synthetic images using 17000 natural images from the PASCAL VOC 2012 [44] dataset
P2	[13]	10k pairs of synthetic images and 110 pairs of real-world images from Zhang et al [7]
P3	[2]	a) Synthetic Reflection Image Dataset (RID) containing 3250 images b) Synthetic dataset using COCO [63]
P4	[4]	a) Synthetic Reflection Image Dataset (RID) b) Synthetic dataset using PASCAL VOC 2012 [44] and COCO [63]
P5	[11]	Synthetic dataset using PLACES 365 [63] and 5552 images from Wan et al [4] (totally 19214 pair of images)
P6	[12]	Synthetic dataset using PASCAL VOC 2012 [44] and SUN2012 [65] datasets
P7	[15]	Synthetic dataset using Open Images [66] dataset
P8	[16]	Synthetic dataset using SUN RGB-D [67] and PASCAL VOC 2012 [44] (10335 images)
P9	[17]	Proposed synthetic dataset using more than 4000 real images
P10	[29]	a) Synthetic dataset from Fan et al [6] b) 200 image pairs (real-world dataset) provided by Li et al [21] called "Nature" c) 90 image pairs provided by Zhang et al [7]
P11	[8]	Two synthetic datasets using PASCAL VOC 2012 [44]
P12	[1]	Synthetic dataset using PASCAL VOC 2012 [66]
P13	[19]	a) Synthetic dataset using random Flickr images b) SIR <sup>2</sup> [38] c) Real-world paired training images from Zhang et al [7].
P14	[9]	a) Synthetic dataset provided by Fan et al [6] b) 90 real-world images provided by Zhang et al [7]
P15	[21]	a) 4k images consisting of 2.8k synthetic images and 1.2k image patches from 290 real-world training images, consisting of 200 images provided by our constructed dataset and 90 images from Berkeley real dataset [7] b) Proposed Nature dataset with 220 real world image pairs
P16	[5]	Synthetic dataset using PASCAL VOC 2012 [66]
P17	[81]	a) 7643 synthetic triplet images using the approach and dataset of PLNet [7] b) 89 real-world image pairs and randomly cropped them into 2000 patches
P18	[23]	a) 537 real images from Wan et al [4] b) Synthetic dataset using ADE20K [68] dataset (30K images) c) Dataset from Zhang et al [7]
P19	[24]	a) 90 real-world images from Berkeley real dataset [7] b) Synthetic dataset from Zhang et al [7] (13700 images)
P20	[25]	a) Synthetic dataset using PASCAL VOC 2012 [44] b) 2k images from Flickr c) 90 real-world images from Zhang et al [7] paper
P21	[26]	Synthesized dataset with DOIDE RGBD [69] and PLACES 365 [64] datasets
P22	[7]	a) 5000 random pairs Flickr b) 90 real-world training images with random resolutions between 256p and 480p
P23	[27]	8.5k synthetic dataset using PASCAL VOC 2012 [44] (6000 for training, 2000 for validation and 500 for test)
P24	[79]	a) 7400 synthetic images using PASCAL-VOC dataset [44] b) Berkeley real dataset [7]
P25	[80]	Synthetic dataset using PLACES 365 [64] dataset and provided dataset by Wan et al. [4]



**TABLE 8. Utilized benchmarking datasets for each selected article.**

ID	Author(s)	Datasets for Experiments (Benchmarking Datasets)
P1	[6]	Synthetic dataset using PASCAL VOC 2012 [44] (850 images)
P2	[13]	a) Test dataset consists of 20 pairs of real images and 140 pairs of synthetic images b) Real images provided by Fan et al [6]
P3	[2]	a) SIR <sup>2</sup> [38] b) The provided validation dataset from the project website of FY17
P4	[4]	SIR <sup>2</sup> [38]
P5	[11]	SIR <sup>2</sup> [38]
P6	[12]	a) SIR <sup>2</sup> [38] b) Real-world images from Zhang et al [7] c) REAL50
P7	[15]	Several real photos
P8	[16]	a) Synthetic dataset using SUN RGB-D [67] and PASCAL VOC 2012 [44] (2860 images) b) 673 real-world images captured by Kinect 2.0.
P9	[17]	SIR <sup>2</sup> [38]
P10	[29]	a) Three datasets from SIR <sup>2</sup> [38] b) Zhang et al [7] c) Nature (Li et al [21])
P11	[8]	Synthetic dataset using PASCAL VOC 2012 [44]
P12	[1]	SIR <sup>2</sup> [38]
P13	[19]	a) SIR <sup>2</sup> [38] b) Real-world paired training data from Zhang et al [7] c) 300 synthetic images d) 25 real-world images for qualitative evaluations
P14	[9]	a) 20 testing real-world images from Zhang et al [7] b) Three sub-dataset from SIR <sup>2</sup> [38]
P15	[21]	a) SIR <sup>2</sup> datasets [38] b) Zhang et al [7] c) The real-world dataset (Nature) [21] with 20 real world image pairs
P16	[5]	a) Real-world images downloaded from Baidu Pictures b) Synthetic dataset using PASCAL VOC 2012 [44]
P17	[81]	a) SIR <sup>2</sup> [38] b) REAL20 [7]
P18	[23]	a) 71 real-world triplets of glass images b) dataset from Fan et al [6] c) Dataset from Zhang et al [7]
P19	[24]	a) Real-world images from PLNet [7] b) SIR <sup>2</sup> [38]
P20	[25]	a) 20 real-world images from PLNet [7] b) SIR <sup>2</sup> [38]
P21	[26]	a) SIR <sup>2</sup> Wild dataset [38] b) Real100 c) 200 rendered images
P22	[7]	Synthetic images from Fan et al [6] dataset and the real images from their real-world dataset
P23	[27]	a) Synthetic images (0.5k images) b) Real-world images from internet
P24	[79]	a) SIR <sup>2</sup> [38] b) Berkeley real dataset [7] c) Their own proposed dataset for evaluation [79]
P25	[80]	a) SIR <sup>2</sup> [38] b) ZC20-ORIEN [82] c) LY20-DATA [21] d) ZN18-DATA [7]



**FIGURE 9.** Percentage of papers conducted experiments on real-world dataset.

#### D. RQ4. WHAT ARE THE ARCHITECTURES OF THE PROPOSED REFLECTION REMOVAL NETWORKS IN EACH PAPER?

This section is merely dedicated to answering the research question RQ4. According to different types of single-image reflection removal networks based on deep learning which are utilized in each publication, we analyzed all the considered research articles with a high focus on the proposed network architectures and deep learning algorithms. In the following, we summarized the proposed deep learning architectures which are utilized for single-image reflection removal.

- In **P1**, Fan *et al.* [6] proposed the CEILNet, which has two cascaded sub-networks: E-CNN which is designed for predicting the edges and I-CNN which is designed for reconstructing the image. In the E-CNN sub-network, they applied a CNN to the source image  $\mathbf{I}^s$  in order to learn an edge map  $\mathbf{E}^l$  of the target image  $\mathbf{I}^l$ . To make the computations easier, they augment the source image  $\mathbf{I}^s$  with the edge map  $\mathbf{E}^s$  as a new input channel. Predicting the edges of the target image is the purpose. The proposed sub-section E-CNN estimates the following function  $f$ . This  $f$  function inputs  $\mathbf{I}^s$  and  $\mathbf{E}^s$  and outputs  $\mathbf{E}^l$ . I-CNN, as the second cascaded sub-network of the CEILNet approximates the following function  $g$  and reconstructs the target image  $\mathbf{I}^l$  by learning the procedure of processing the input image  $\mathbf{I}^s$  given the target image map  $\mathbf{E}^l$  which the first cascaded sub-network (E-CNN) predicts the  $g$  function. This  $g$  function inputs  $\mathbf{I}^s$  and  $\mathbf{E}^l$ , and then outputs  $\mathbf{I}^l$  [6].
- In **P2**, Fu *et al.* [13] proposed an algorithm that solves the reflection removal issue in two phases. In the first phase, they used an instance segmentation network to obtain the region-aware map  $\mathbf{M}'$  from the input image. In the first phase, they used the Mask R-CNN network [70]. In the second phase, they input both the original image and also  $\mathbf{M}'$  together into a CNN which performs as an encoder-decoder in order to erase the reflection  $\mathbf{B}'$  from the original image [13].
- In **P3**, Wan *et al.* [2] proposed CoRRN which contains three sub-networks for estimation of the background and reflections in a cooperative way. These three sub-networks are CencN, idecN and GdecN. Context encoder network (CencN) is designed based on VGG16 [37] and is responsible for suppressing the sparse reflection residues and then extracting information related to the scenes from different layers of images. Another responsibility of CencN is facilitating the training networks. Gradient decoder network (GdecN) is also responsible for learning a mapping from  $\mathbf{I}$  to  $\nabla\mathbf{B}$ . and finally, image decoder network (IdecN) which is a multi-tasking learning network is responsible for learning a mapping from  $\mathbf{I}$  to  $\mathbf{B}$  and  $\mathbf{R}$ . In the framework of CoRRN, the related gradient features from GdecN sub-network guides the IdecN sub-network [2].
- In **P4**, Wan *et al.* [4] proposed the CRRN with a multi-task learning approach that evaluates the background and the reflection of the image using  $\nabla\mathbf{B}$ . The proposed framework of the CRRN contains two cooperative blocks: first, the gradient interference network (GiN) which is designed in order to estimate and evaluate gradients of the background image and second, the image interference network (IiN) which is designed to estimate and evaluate the background and reflection layers of the image. The GiN is fed with the blend image and corresponding gradients as a tensor with four channels. IiN is also fed with a mixture of images that consists of reflections. The phase of up-sampling in the IiN is guided by some gradient features that are provided in GiN with the same resolution. The IiN block has two feature extraction layers for extracting the features of scale-invariant correlated with the background image. The output of the IiN block will be the estimated background and reflection image and the output of the GiN block will be the estimated gradient of the background image [4].
- In **P5**, Zheng *et al.* [11] proposed a network that aims to remove droplets of rain and reflections in the glass or transparent material from a single-image. The two sub-networks  $\mathbf{g}$  and  $\mathbf{h}$  are both created of two CNN layers with the size of two for the process of down-sampling, 6 residual units [31] and also two interpolating convolutional neural layers with the size of two for the process of up-sampling that has instance normalization [32]. Each of the  $\mathbf{g}$  and  $\mathbf{h}$  sub-networks consists of 1 and 3 output channels [11].
- In **P6**, Zhang *et al.* [12] proposed an algorithm that contains two cascaded subnetworks: decomposition network (D-Net) and refinement network (R-Net).

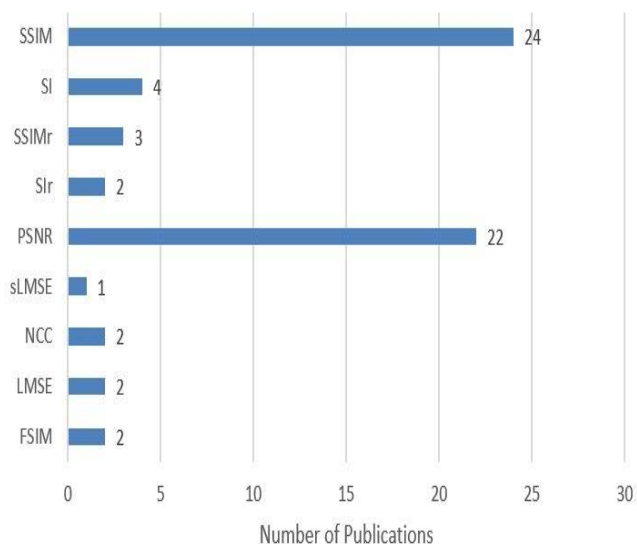
- The D-Net sub-network is decomposing the input image which is called  $\mathbf{I}$  into the background layer and the reflection layer. Then the R-Net sub-network refines the output image. This procedure in the R-Net is done by adding back the details that are ignored by the D-Net sub-network with the help of an attention strategy. Also, Zhang *et al.* add a former sub-network that involves human interaction into the proposed cascaded network. They request the user to show the reflection and background areas in the considered image. The guidance input of the user is then transformed to a pyramid of edge maps to recognize the difference between the reflection and background areas of the image [12].
- In **P7**, Chou *et al.* [15] proposed a network which is a GAN with two generator blocks to model the reflection areas separately that has different intensity [15].
  - In **P8**, Chang *et al.* [16] proposed a network that consists of 3 sub-networks. They used the first four units of VGG-19 as the encoder. The RNN and DEN are the two sub-networks that are designed as decoders. The image with reflection  $\mathbf{I}$  is given to the encoder and the output of this sub-network is given to RRN and DEN subnetworks. RRN is the proposed sub-network for the reflection removal task that outputs the evaluated transmission layer which is called  $\mathbf{O}$ . The evaluated reflection  $\mathbf{R}'$  is obtained by  $\mathbf{I}-\mathbf{O}$ . And DEN sub-network is designed for depth estimation which has the transmission depth  $\mathbf{D}_T$  as the output.  $\mathbf{rD}_T$  is also the refined depth map [16].
  - In **P9**, Ma *et al.* [17] proposed a network that consists of two mapping functions. Generator  $\mathbf{G}: (\mathbf{B}, \mathbf{R}) \rightarrow \mathbf{M}$  and separator  $\mathbf{S}: \mathbf{M} \rightarrow (\mathbf{B}, \mathbf{R}, \mathbf{E})$ , which  $\mathbf{M}$  is a real-world mixed image,  $\mathbf{B}$  is background and  $\mathbf{R}$  is reflection. Generator  $\mathbf{G}$  and the separator  $\mathbf{S}$  both share similar architectures. They both have a down-sampling unit with two CNN layers for increasing the receptive field size and a feature extraction unit that has 9 residual blocks for the process of extracting robustly and an up-sampling unit that consists of two transposed CNN layers. The generator  $\mathbf{G}$  consists of two down-sampling units in order to get the input of reflection and background. The separator  $\mathbf{S}$  also consists of a multi-tasking unit and three up-sampling units to make more accurate results. They have utilized  $70 \times 70$  PatchGANs [71], [72] in the discriminator unit to apply the algorithm in arbitrary-sized images [17].
  - In **P10**, Dong *et al.* [29] proposed an algorithm that consists of two phases. In the first phase, the algorithm predicts the RCMAP and the reflection layer. In the second phase, the algorithm predicts the transmission layer CBAM [29].
  - In **P11**, Yang *et al.* [8] proposed BDN network architecture that has three sub-networks  $G_0(\cdot)$ ,  $H(\cdot)$  and  $G_1(\cdot)$ . Ronneberge *et al.* used a kind of U-net [33], [34] which has an encoder and a decoder part for the implementation of these three sub-networks. In this network architecture, the second and third sub-networks contain the same network structure but with different parameters. The first sub-network has 14 convolutional layers, in contrast, the second and third sub-networks have 10 convolutional layers. Note that  $G_0(\cdot)$ ,  $H(\cdot)$  and  $G_1(\cdot)$  are all cascaded sub-networks [8].
  - In **P12**, Abiko and Ikehara [1] proposed an algorithm based on GAN with gradient constraint. This algorithm has three sub-networks Generator, discriminator, and feature extractor that are operating based on CNN. The proposed sub-network generator is performing based on UNet++L<sup>4</sup> [35] that is a mixture of CNN layers, leaky ReLU layers, max-pooling layers, batch normalization layers and bilinear interpolation layers. Due to the fact that the authors implement a deep supervision network, four outputs are available in this network. In this paper, Abiko and Ikehara [1] utilized  $\hat{\mathbf{B}}$  for the main output of the proposed network and the other three outputs as a tool for computing pixel loss. The discriminator block includes some CNNs, batch normalization and leaky ReLU layer. And eventually, they used the L<sub>2</sub> difference for the output of the discriminator block which has a size of  $16 \times 16$  in order to calculate the adversarial loss [1].
  - In **P13**, Wen *et al.* [19] proposed SynNet which is operating based on an encoder-decoder architecture and has six channels for image input. The encoder and decoder, both consist of three CNN layers, and between these two encoders and decoders, nine residual blocks are added to improve the illustration of reflection properties. InstanceNorm layer [36] and ReLU activation function are included in all CNN layers except the last one which is followed by the Sigmoid activation function in order to scale the output. Eventually, the proposed network contains a three-channel output with an alpha blending mask [19].
  - In **P14**, Wei *et al.* [9] proposed a network for single-image reflection removal. They trained a feed-forward CNN  $G_{\theta G}$ . In the first step, a VGG-19 [37] sub-network is used. Then the output of VGG-19 is up-sampled and fed to the CNN  $G_{\theta G}$  network. The CNN  $G_{\theta G}$  contains 3 layers of Conv-ReLU pairs, 13 residual blocks, Conv-ReLU pair and Pyramid pooling [9].
  - In **P15**, Li *et al.* [21] proposed IBCLN which contains two blocks: a transmission-prediction block  $G_T$  and a reflection-prediction block  $G_R$ . Both of these two proposed blocks are designed based on convolutional LSTM networks.  $G_T$  block learns the transmission  $\mathbf{T}$  and  $G_R$  tries to learn the residual reflection  $\tilde{\mathbf{R}}$ . Both block use an encoder that contains 11 layers of convolutional LSTM layers for extracting the features of the input, and a decoder that contains 8 CNN layers that generate the predicted residual reflection layer of the transmission layer. The LSTM layers have a Sigmoid or Tanh activation function. All of the other convolutional layers are followed by a ReLU activation function [21].

- In **P16**, Chang and Jung [5] proposed a network based on an encoder-decoder architecture for reflection removal tasks. The encoder sub-network **E** learns dimensional features via CNN layers and the point-reversed procedure begins to operate. The decoder sub-network is connected to the **E** sub-network in order to map the high dimensional features reversely to natural images using some deconvolutions. Also, three skips are appended between **E** and **D** sub-networks to save the resolution of the outputs [5].
- In **P17** [81], the proposed network first separates the extracted features into two branches using feature learning: background constituent and reflection constituent. The contrastive feature decomposition sub-network achieves more accurate feature decomposition by their proposed contrastive supervision algorithm. In the end, the dense feature refinement sub-network tries to refine the details of restored images in order to accomplish images with high standard including both the background and the reflection images [81].
- In **P18**, Han and Sim [23] proposed a reflection removal architecture that contains an interpreter and three multi-scale generators. All of the designed generators predict both transmission image  $T_k$  and also the image with reflection  $R_k$ . Then they are up-sampled and concatenated to the input image to become the input image of the generator at the better image scale  $k+1$ . In the proposed architecture, three generators are included with the exact same structure. These generators supply the product images at the scales of 0.25, 0.5 and 1.0 in comparison to the primary input image with reflections. The proposed network also utilizes an encoder-decoder algorithm, and two CNN layers followed by ReLU and batch normalization activation functions [23].
- In **P19**, Cheng *et al.* [24] presented a network that contains two parts: a generator and a 5-layer discriminator. The generative sub-network also contains two sub-parts: an encoder-decoder block and a FCN block. The discriminator is operating with SNGAN loss [24].
- In **P20**, Chang *et al.* [25] proposed a single image reflection which at first, edge estimator  $\epsilon$  is being trained. In this process, it takes the reflection contaminated image **I** as input and estimates the edge map of the transmission layer **T**. Then, reflection classifier **C** is being trained in order to be used to recognize whether the reflection is available or not. The classifier also takes the reflection layer **R** and concatenates it with the reflection contaminated image **I** or the transmission layer **T** as input pair and eventually outputs the matching label of the pair. In other words, the proposed network takes **I** as the input of the architecture, then it produces the estimated transmission and then in the second stage, obtains the ultimate prediction of the transmission layer by recurrent decomposition [25].
- The proposed network architecture in **P21** [26] has two sub-networks: SP-net, a major separation sub-network and BT-net, a backtrack network for the back scene reflection. The SP-net separates the input image (image with a reflection into  $T^*$  and  $R^*$  (with glass-effect). In the second phase,  $R^*$  is given to the BT-net in order to erase glass/lens effect such as ghosting, distortion, defocusing and attenuation.  $R^*$  is the output of the BT-net.  $R^*$  must be glass/lens effect-free.  $R^*$  is also utilized to provide more error computation for SP-net. All inputs of the proposed sub-networks are appended with extracted features from the famous VGG-19 [37] network [26].
- In **P22**, Zhang *et al.* [7] proposed a single-image reflection separation network. To train the network using a semantic understanding of the input image, they defined hyper-column features [74] by extracting these features using VGG-19 [37] network. This VGG-19 was pre-trained on the ImageNet dataset [75]. And finally, the proposed network **f** is a FCN that is so much resembling to context aggregation networks [76], [77]. In other words, the network is using a FCN trained thoroughly with losses that make use of both low and high level information available in the image [7].
- In **P23**, Heo and Choe [27] proposed a single-image reflection removal that operates using conditional GANs. The proposed reflection removal network in **P23** [27] mainly adopts the basic network of pix2pix [34] and uses the ADAM optimizer with a learning rate 0.0002. The Generator block uses the proposed U-Net [33] and the discriminator block employs a classifier that operates based on FCNs. The ADAM optimizer [73] is also applied with a learning rate  $2 \times 10^{-4}$ . The momentum parameters are  $\beta_1 = 0.5$  and  $\beta_2 = 0.999$  [27].
- In **P24**, an image of resolution  $H \times W$  with reflection is given to the network as input. The proposed network determines the number of scales  $N$ . In the next stage, a  $N$ -scale space representation of the input image is constructed from the scale space in two phases: Low scale sub-network (LSSNet) and b) Progressive Inference (PI) of the higher scales using Convolutional Guided Filter up-sampling (CGF) and High Scale Sub-network (HSSNet) [79].
- In **P25**, **I** and **OT** are given to the proposed sub-network **g** as inputs and then the sub-network outputs  $e_{pre}$  and  $e'_{pre}$ . In the second phase, the proposed sub-network **h** takes the input which is the concatenation of **I** and spatially-replicated  $e_{pre}$ , and gives  $T_{pre}$  as the output [80].

#### **E. RQ5. WHAT QUANTITATIVE METRICS HAVE BEEN USED TO ACCREDIT THE EFFICIENCY OF THE PROPOSED DEEP LEARNING TECHNIQUES?**

To answer RQ5, we gathered all the relevant information about the used quantitative metrics in each selected research article. Many evaluation metrics were utilized in research articles to assess and evaluate the performance of the

proposed methods. Figure 10 presents the number of studies using different quantitative metrics including FSIM, LMSE, NCC, sLMSE, PSNR, SI<sub>r</sub>, SSIM<sub>r</sub>, SI, and SSIM to evaluate their proposed method.



**FIGURE 10.** Number of research papers using FSIM, LMSE, NCC, sLMSE, PSNR, SI<sub>r</sub>, SSIM<sub>r</sub>, SI and SSIM as a quantitative metric for evaluation of their proposed deep learning technique.

All of the papers used multiple quantitative metrics for performance validation, reporting the results of the proposed method and comparing their method with other research works. According to figure 10, almost 96% of the selected research articles used SSIM quantitative metric in order to validate the performance of their research work. In the second stage, near 88% of the selected research articles used PSNR as a quantitative metric. Therefore, it is obvious that SSIM and PSNR are the most utilized quantitative metrics in the selected research articles. Among these 25 articles, 16% of the selected articles used SI metric, 12% used SSIM<sub>r</sub>, 8% used FSIM, LMSE, NCC, SI<sub>r</sub>, and at the end, 4% used sLMSE. Due to the fact that SSIM and PSNR metrics are broadly popular in the research field, they can be used to compare the results of each newly proposed technique with previous research works. Finally, in Table 9, we presented more details about the results of our analysis over the metrics for each paper.

**F. RQ6. WHAT ARE THE CURRENT LIMITATIONS AND CHALLENGES IN THIS RESEARCH FIELD?**

In this section, we summarize the weakness and limitations represented by the deep learning-based single-image reflection removal algorithms as recorded by researchers to be resolved in the future of nominated research works. In this SLR, we have extracted all the relevant information about the limitations, challenges and weaknesses from the limitations section of the considered research articles. According to the

**TABLE 9.** Analysis of each selected paper according to the utilized quantitative metrics for performance validation.

ID	SSIM	SI	SSIM <sub>r</sub>	SI <sub>r</sub>	PSNR	sLMSE	NCC	FSIM	LMSE
P1	✓	✗	✗	✗	✓	✗	✗	✗	✗
P2	✓	✗	✗	✗	✓	✗	✗	✗	✗
P3	✓	✓	✓	✓	✓	✗	✗	✗	✗
P4	✓	✓	✓	✓	✗	✗	✗	✗	✗
P5	✓	✓	✗	✗	✓	✗	✗	✗	✗
P6	✓	✓	✗	✗	✗	✓	✓	✗	✗
P7	✗	✗	✗	✗	✗	✗	✗	✗	✗
P8	✓	✗	✗	✗	✓	✗	✗	✓	✗
P9	✓	✗	✓	✗	✓	✗	✗	✗	✗
P10	✓	✗	✗	✗	✓	✗	✗	✗	✗
P11	✓	✗	✗	✗	✓	✗	✗	✗	✗
P12	✓	✗	✗	✗	✓	✗	✗	✗	✗
P13	✓	✗	✗	✗	✓	✗	✗	✗	✗
P14	✓	✗	✗	✗	✓	✗	✓	✗	✓
P15	✓	✗	✗	✗	✓	✗	✗	✗	✗
P16	✓	✗	✗	✗	✓	✗	✗	✓	✗
P17	✓	✗	✗	✗	✓	✗	✗	✗	✗
P18	✓	✗	✗	✗	✓	✗	✗	✗	✓
P19	✓	✗	✗	✗	✓	✗	✗	✗	✗
P20	✓	✗	✗	✗	✓	✗	✗	✗	✗
P21	✓	✗	✗	✗	✓	✗	✗	✗	✗
P22	✓	✗	✗	✗	✓	✗	✗	✗	✗
P23	✓	✗	✗	✗	✓	✗	✗	✗	✗
P24	✓	✗	✗	✗	✓	✗	✗	✗	✗
P25	✓	✗	✗	✗	✓	✗	✗	✗	✗

selected papers, the limitations and challenges are explained as follows:

- It is possible to encounter an image inpainting problem that is more challenging and more complicated because of the loss of the background information in some parts of the image that contains saturated reflections. Due to this problem, almost all of the proposed techniques cannot perform effectively in order to remove the reflections from the image. [2].
- With the existence of many convolutional layers in the proposed networks, the color-shifting problems are one of the challenges for the proposed reflection removal networks [2], [4].



- Some of the proposed networks did not consider outlooks that are vertical to the glass in their proposed algorithm and also they did not take into account curved glass or glass with a particular shape. In some other networks, the diversity of scenarios and capturing settings for the images which are included in the synthetic dataset needs to be improved. These problems in data generation may restrict the generalization ability of the dataset [2], [26]. The synthesized and real-world images in some datasets of the proposed algorithms are all indoor scenes that may not be suitable for outdoor scenes. These proposed networks were not very effective and almost fail in outdoor scenes [16].
- When the whole images are dominated by reflection or ghosting reflection which makes it so hazy and blurry, or the reflection layers and background are overlapped, the effectiveness of the proposed networks may drop and these networks may not be able to completely remove the reflections and the evaluated background still remains visible residual edges. Also, the proposed technique may have some problems with gradient disappearing when the deep learning technique is trained directly on the images [4], [12], [16].
- According to the fact that some of the presented networks are operating based on the extracted edges, these algorithms may not work properly whenever there is a loss of edge information or the edge information is low-confident [12].
- Some of the proposed single-image reflection removal networks have some limitations in the process of performing on real-world images. This limitation happens with a lack of a real-world training dataset [27].

#### **G. RQ7. WHAT ARE THE POSSIBLE FUTURE WORKS AND DIRECTIONS IN THE SINGLE-IMAGE REFLECTION REMOVAL FIELD VIA DEEP LEARNING TECHNIQUES?**

Due to the current bibliographic review, we encounter the potential directions and future work recommended by researchers as mentioned below:

- Applying the proposed networks on more image processing and image layer decomposition projects. Outstanding outcomes for image denoising, inpainting, raindrop removal, flare removal and dehazing have been reached in some introductory investigations and experiments [6], [11], [21].
- Improving the proposed networks in some cases in which the whole image is dominated by reflections in order to have better results. Continuing the work on the proposed networks to enhance the capability for dealing with challenging images with reflections [4], [7], [11], [15], [24].
- Some of the proposed networks rely on hand-crafted features. Proposing and designing a more hand-free and automated reflection removal algorithm than the proposed ones which can free users from guidance and suppress reflection with high quality can be mentioned as a future direction and it is expected to successfully deal with the limitations in challenging reflection removal tasks [12], [23].
- The dataset used for training the proposed networks is mostly built based on the screen blending method as the researchers assert that it is more precise to depict the real reflection than the mostly utilized summative method [15].
- Simplifying the proposed networks to reduce the number of existing parameters and upgrade the presumption rapidly for its usages on the mobile computing programs [29], [80].
- Different kinds of networks including generative models inspired by GANs can be applied to this single-image reflection removal task to achieve better results. Moreover, due to the uncertain nature of the reflection layer, generative models based on an uncertainty map can be used in future work [27].

#### **IV. CONCLUSION**

In this systematic review, we aimed to explore and analyze the academic research papers in the field of single-image reflection removal using deep learning and neural network techniques. For this purpose, we performed a complete and comprehensive statistical analysis on this topic by extracting the answers to 7 research questions which are provided below:

- RQ1: What is the distribution of the selected research articles over the last seven years and their types?
- RQ2: What datasets/databases are used to train the proposed deep learning model in considered publications?
- RQ3: What datasets/databases are used for conducting experiments on single-image reflection removal?
- RQ4: What are the architectures of the proposed reflection removal networks in each paper?
- RQ5: What quantitative metrics have been used to accredit the efficiency of the proposed deep learning techniques?
- RQ6: What are the current limitations and challenges in this research field?
- RQ7: What are the possible future works and directions in the single-image reflection removal field via deep learning techniques?

By the use of our research methodology, we investigated almost 1600 research articles published between the years 2015 and 2021. Twenty-five of them have been selected in this SLR to be discussed. In order to answer these proposed questions, our research team identified the distribution of the selected articles (RQ1), datasets and databases used for training, testing and experiments (RQ2 and RQ3), structures and architectures of the proposed algorithms (RQ4), quantitative metrics for performance validation (RQ5) and finally, limitations, challenges and future work of each research studies (RQ6 and RQ7).

In the process of conducting this SLR, we observed that the current trend in the reflection removal field is to use

convolutional neural networks and deep learning techniques. The advances in this field are mostly related to using different CNN structures and architectures and also using datasets with more images. One of the expected results of this SLR is that in spite of the very advanced deep learning algorithms utilized for this reflection problem, there is still a lot of work left to do in this research field in the future. We note that it is discussed in detail in the RQ.7 about the possible future works. Another observation is that there are few studies for single-image reflection removal using deep learning (25 studies were identified). We also observed that there is a lack of real-world training and benchmark single-image reflection removal datasets in order to train or evaluate the deep learning algorithms. Further studies can be on making a comprehensive training and benchmark dataset.

In conclusion, our research team hopes that due to the quite wide topic coverage, this systematic review can be beneficial and helpful for the many AI researchers and engineers who are beginning research in the field of single-image reflection removal and deep learning. In spite of the fact that reflection removal using deep learning techniques has newly been a trend among researchers, we strongly believe that fast progress and advances in this field can be very important to reach a higher level of services in computer vision and image processing applications.

## ACKNOWLEDGMENT

The authors would like to thank Ramin Abolfazl Soltani, a Research Assistant at the Laboratory of Movement Analysis and Measurement at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He accompanied their research team in the procedure of reviewing this article. They also shared long sessions with him about the process of conducting this systematic literature review.

## REFERENCES

- [1] R. Abiko and M. Ikehara, "Single image reflection removal based on GAN with gradient constraint," *IEEE Access*, vol. 7, pp. 148790–148799, 2019, doi: [10.1109/ACCESS.2019.2947266](https://doi.org/10.1109/ACCESS.2019.2947266).
- [2] R. Wan, B. Shi, H. Li, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "CoRRN: Cooperative reflection removal network," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 12, pp. 2969–2982, Dec. 2020, doi: [10.1109/TPAMI.2019.2921574](https://doi.org/10.1109/TPAMI.2019.2921574).
- [3] J. Sun, Y. Chang, C. Jung, and J. Feng, "Multi-modal reflection removal using convolutional neural networks," *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 1011–1015, Jul. 2019, doi: [10.1109/LSP.2019.2915560](https://doi.org/10.1109/LSP.2019.2915560).
- [4] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "CoRRN: Multi-scale guided concurrent reflection removal network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 4777–4785.
- [5] Y. Chang and C. Jung, "Single image reflection removal using convolutional neural networks," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1954–1966, Apr. 2019, doi: [10.1109/TIP.2018.2880088](https://doi.org/10.1109/TIP.2018.2880088).
- [6] Q. Fan, J. Yang, G. Hua, B. Chen, and D. Wipf, "A generic deep architecture for single image reflection removal and image smoothing," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 3238–3247.
- [7] X. Zhang, R. Ng, and Q. Chen, "Single image reflection separation with perceptual losses," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4786–4794.
- [8] J. Yang, D. Gong, L. Liu, and Q. Shi, "Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 654–669.
- [9] K. Wei, J. Yang, Y. Fu, D. Wipf, and H. Huang, "Single image reflection removal exploiting misaligned training data and network enhancements," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8178–8187.
- [10] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," *Engineering*, vol. 45, no. 4, p. 1051, 2007.
- [11] Q. Zheng, B. Shi, X. Jiang, L.-Y. Duan, and A. C. Kot, "Denoising adversarial networks for rain removal and reflection removal," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2766–2770, doi: [10.1109/ICIP.2019.8803225](https://doi.org/10.1109/ICIP.2019.8803225).
- [12] H. Zhang, X. Xu, H. He, S. He, G. Han, J. Qin, and D. Wu, "Fast user-guided single image reflection removal via edge-aware cascaded networks," *IEEE Trans. Multimedia*, vol. 22, no. 8, pp. 2012–2023, Aug. 2020, doi: [10.1109/TMM.2019.2951461](https://doi.org/10.1109/TMM.2019.2951461).
- [13] R. Fu, P. Kuang, Y. Zhou, H.-R. Yan, and T.-Y. Zheng, "Area-aware reflection detection and removal for single image," in *Proc. 16th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process.*, Dec. 2019, pp. 307–310, doi: [10.1109/ICCWAMTIP47768.2019.9067633](https://doi.org/10.1109/ICCWAMTIP47768.2019.9067633).
- [14] T. Li, Y.-H. Chan, and D. P. K. Lun, "Improved multiple-image-based reflection removal algorithm using deep neural networks," *IEEE Trans. Image Process.*, vol. 30, pp. 68–79, 2021, doi: [10.1109/TIP.2020.3031184](https://doi.org/10.1109/TIP.2020.3031184).
- [15] N.-H. Chou, L.-C. Chuang, and M.-S. Lee, "Intensity-aware GAN for single image reflection removal," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2019, pp. 590–594, doi: [10.1109/APSIPAASC47483.2019.9023087](https://doi.org/10.1109/APSIPAASC47483.2019.9023087).
- [16] Y. Chang, C. Jung, and J. Sun, "Joint reflection removal and depth estimation from a single image," *IEEE Trans. Cybern.*, vol. 51, no. 12, pp. 5836–5849, Dec. 2021, doi: [10.1109/TCYB.2019.2959381](https://doi.org/10.1109/TCYB.2019.2959381).
- [17] D. Ma, R. Wan, B. Shi, A. Kot, and L. Duan, "Learning to jointly generate and separate reflections," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2444–2452, doi: [10.1109/ICCV.2019.00253](https://doi.org/10.1109/ICCV.2019.00253).
- [18] B. Rупing, L. Jing, Y. Tao, L. Zhaoyang, L. Zhongzhen, and Z. Fangbing, "Removing reflection from images based on BP neural network of motion field," in *Proc. IEEE 13th Int. Conf. Signal Process. (ICSP)*, Nov. 2016, pp. 768–771, doi: [10.1109/ICSP.2016.7877935](https://doi.org/10.1109/ICSP.2016.7877935).
- [19] Q. Wen, Y. Tan, J. Qin, W. Liu, G. Han, and S. He, "Single image reflection removal beyond linearity," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3766–3774, doi: [10.1109/CVPR.2019.00389](https://doi.org/10.1109/CVPR.2019.00389).
- [20] T. Takahashi, K. Uruma, and K. Kobayashi, "Single image reflection removal for privacy protection using deep CNN," in *Proc. IEEE 63rd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2020, pp. 1028–1031, doi: [10.1109/MWSCAS48704.2020.9184486](https://doi.org/10.1109/MWSCAS48704.2020.9184486).
- [21] C. Li, Y. Yang, K. He, S. Lin, and J. E. Hopcroft, "Single image reflection removal through cascaded refinement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3562–3571, doi: [10.1109/CVPR42600.2020.00362](https://doi.org/10.1109/CVPR42600.2020.00362).
- [22] Z. Chi, X. Wu, X. Shu, and J. Gu, "Single image reflection removal using deep encoder-decoder network," 2018, *arXiv:1802.00094*.
- [23] B.-J. Han and J.-Y. Sim, "Single image reflection removal using non-linearly synthesized glass images and semantic context," *IEEE Access*, vol. 7, pp. 170796–170806, 2019, doi: [10.1109/ACCESS.2019.2955994](https://doi.org/10.1109/ACCESS.2019.2955994).
- [24] K. Cheng, J. Song, J. Du, S. Rong, and H. Zhou, "Single image reflection removal via attention model and SN-GAN," *IEEE Access*, vol. 8, pp. 96046–96054, 2020, doi: [10.1109/ACCESS.2020.2995871](https://doi.org/10.1109/ACCESS.2020.2995871).
- [25] Y.-C. Chang, C.-N. Lu, C.-C. Cheng, and W.-C. Chiu, "Single image reflection removal with edge guidance, reflection classifier, and recurrent decomposition," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 2032–2041, doi: [10.1109/WACV48630.2021.00208](https://doi.org/10.1109/WACV48630.2021.00208).
- [26] S. Kim, Y. Huo, and S.-E. Yoon, "Single image reflection removal with physically-based training images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5163–5172, doi: [10.1109/CVPR42600.2020.00521](https://doi.org/10.1109/CVPR42600.2020.00521).
- [27] M. Heo and Y. Choe, "Single-image reflection removal using conditional GANs," in *Proc. Int. Conf. Electron., Inf., Commun. (ICEIC)*, Jan. 2019, pp. 1–4, doi: [10.23919/ELINFOCOM.2019.8706433](https://doi.org/10.23919/ELINFOCOM.2019.8706433).
- [28] T. Li and D. P. K. Lun, "Single-image reflection removal via a two-stage background recovery process," *IEEE Signal Process. Lett.*, vol. 26, no. 8, pp. 1237–1241, Aug. 2019, doi: [10.1109/LSP.2019.2926828](https://doi.org/10.1109/LSP.2019.2926828).
- [29] Z. Dong, K. Xu, Y. Yang, H. Bao, W. Xu, and R. W. H. Lau, "Location-aware single image reflection removal," 2020, *arXiv:2012.07131*.

- [30] Y. Li and M. S. Brown, "Exploiting reflection change for automatic reflection removal," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2432–2439.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [32] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*.
- [33] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI (Lecture Notes in Computer Science)*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.
- [34] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.
- [35] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham, Switzerland: Springer, 2018, pp. 3–11.
- [36] D. Ulyanov, V. Lebedev, A. Vedaldi, and V. S. Lempitsky, "Texture networks: Feed-forward synthesis of textures and stylized images," in *Proc. ICML*, 2016, pp. 1349–1357.
- [37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [38] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "Benchmarking single-image reflection removal algorithms," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 3942–3950, doi: [10.1109/ICCV.2017.423](https://doi.org/10.1109/ICCV.2017.423).
- [39] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1647–1654, Sep. 2007, doi: [10.1109/TPAMI.2007.1106](https://doi.org/10.1109/TPAMI.2007.1106).
- [40] Y. Shih, D. Krishnan, F. Durand, and W. T. Freeman, "Reflection removal using ghosting cues," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3193–3201, doi: [10.1109/CVPR.2015.7298939](https://doi.org/10.1109/CVPR.2015.7298939).
- [41] R. Wan, B. Shi, L. Duan, A. Tan, W. Gao, and A. C. Kot, "Region-aware reflection removal with unified content and gradient priors," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2927–2941, Jun. 2018, doi: [10.1109/TIP.2018.2808768](https://doi.org/10.1109/TIP.2018.2808768).
- [42] R. Wan, B. Shi, H. Li, L.-Y. Duan, and A. C. Kot, "Reflection scene separation from a single image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2395–2403, doi: [10.1109/CVPR42600.2020.00247](https://doi.org/10.1109/CVPR42600.2020.00247).
- [43] 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.
- [44] M. Al-Sarem, W. Boulila, M. Al-Harby, J. Qadir, and A. Alsaedi, "Deep learning-based rumor detection on microblogging platforms: A systematic review," *IEEE Access*, vol. 7, pp. 152788–152812, 2019, doi: [10.1109/ACCESS.2019.2947855](https://doi.org/10.1109/ACCESS.2019.2947855).
- [45] J. Memon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR)," *IEEE Access*, vol. 8, pp. 142642–142668, 2020, doi: [10.1109/ACCESS.2020.3012542](https://doi.org/10.1109/ACCESS.2020.3012542).
- [46] M. A.-M. Khan, M. R. J. Khan, A. Tooshil, N. Sikder, M. A. P. Mahmud, A. Z. Kouzani, and A.-A. Nahid, "A systematic review on reinforcement learning-based robotics within the last decade," *IEEE Access*, vol. 8, pp. 176598–176623, 2020, doi: [10.1109/ACCESS.2020.3027152](https://doi.org/10.1109/ACCESS.2020.3027152).
- [47] A. W. Li and G. S. Bastos, "Stock market forecasting using deep learning and technical analysis: A systematic review," *IEEE Access*, vol. 8, pp. 185232–185242, 2020, doi: [10.1109/ACCESS.2020.3030226](https://doi.org/10.1109/ACCESS.2020.3030226).
- [48] A. Saboor, T. Kask, A. Kuusik, M. M. Alam, Y. Le Moulec, I. K. Niazi, A. Zoha, and R. Ahmad, "Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review," *IEEE Access*, vol. 8, pp. 167830–167864, 2020, doi: [10.1109/ACCESS.2020.3022818](https://doi.org/10.1109/ACCESS.2020.3022818).
- [49] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019, doi: [10.1109/ACCESS.2019.2896880](https://doi.org/10.1109/ACCESS.2019.2896880).
- [50] M. A. Kuhail, S. Farooq, R. Hammad, and M. Bahja, "Characterizing visual programming approaches for end-user developers: A systematic review," *IEEE Access*, vol. 9, pp. 14181–14202, 2021, doi: [10.1109/ACCESS.2021.3051043](https://doi.org/10.1109/ACCESS.2021.3051043).
- [51] T.-S. Reinel, R.-P. Raul, and I. Gustavo, "Deep learning applied to steganalysis of digital images: A systematic review," *IEEE Access*, vol. 7, pp. 68970–68990, 2019, doi: [10.1109/ACCESS.2019.2918086](https://doi.org/10.1109/ACCESS.2019.2918086).
- [52] A. Bajaj and O. P. Sangwan, "A systematic literature review of test case prioritization using genetic algorithms," *IEEE Access*, vol. 7, pp. 126355–126375, 2019, doi: [10.1109/ACCESS.2019.2938260](https://doi.org/10.1109/ACCESS.2019.2938260).
- [53] S. I. Hassan, M. M. Alam, U. Illahi, M. A. Al Ghamdi, S. H. Almotiri, and M. M. Su'ud, "A systematic review on monitoring and advanced control strategies in smart agriculture," *IEEE Access*, vol. 9, pp. 32517–32548, 2021, doi: [10.1109/ACCESS.2021.3057865](https://doi.org/10.1109/ACCESS.2021.3057865).
- [54] N. Rajabli, F. Flammini, R. Nardone, and V. Vittorini, "Software verification and validation of safe autonomous cars: A systematic literature review," *IEEE Access*, vol. 9, pp. 4797–4819, 2021, doi: [10.1109/ACCESS.2020.3048047](https://doi.org/10.1109/ACCESS.2020.3048047).
- [55] A. Baldominos, A. Puello, H. Oğul, T. Aşuroğlu, and R. Colomo-Palacios, "Predicting infections using computational intelligence—A systematic review," *IEEE Access*, vol. 8, pp. 31083–31102, 2020, doi: [10.1109/ACCESS.2020.2973006](https://doi.org/10.1109/ACCESS.2020.2973006).
- [56] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman, "Ground truth dataset and baseline evaluations for intrinsic image algorithms," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2335–2342.
- [57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [58] T. Xue, M. Rubinstein, C. Liu, and W. T. Freeman, "A computational approach for obstruction-free photography," *ACM Trans. Graph.*, vol. 34, no. 4, p. 79, Jul. 2015.
- [59] B. Kitchenham, R. Pretorius, D. Budgen, O. P. Brereton, M. Turner, M. Niazi, and S. Linkman, "Systematic literature reviews in software engineering—A tertiary study," *Inf. Softw. Technol.*, vol. 52, no. 8, pp. 792–805, 2010.
- [60] Y. Li and M. S. Brown, "Single image layer separation using relative smoothness," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2752–2759.
- [61] Y. Diamant and Y. Y. Schechner, "Overcoming visual reverberations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [62] X. Guo, X. Cao, and Y. Ma, "Robust separation of reflection from multiple images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2187–2194.
- [63] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 740–755.
- [64] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1452–1464, Jun. 2018.
- [65] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, "SUN database: Large-scale scene recognition from abbey to zoo," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3485–3492.
- [66] I. Krasin, T. Duerig, N. Alldrin, A. Veit, S. Abu-El-Haija, S. Belongie, D. Cai, Z. Feng, V. Ferrari, V. Gomes, A. Gupta, D. Narayanan, C. Sun, G. Chechik, and K. Murphy. (2016). *OpenImages: A Public Dataset for Large-Scale Multi-Label and Multi-Class Image Classification*. [Online]. Available: <https://github.com/openimages>
- [67] S. Song, S. P. Lichtenberg, and J. Xiao, "SUN RGB-D: A RGB-D scene understanding benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 567–576.
- [68] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, "Scene parsing through ADE20K dataset," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5122–5130.
- [69] I. Vasiljevic, N. Kolkin, S. Zhang, R. Luo, H. Wang, F. Z. Dai, A. F. Daniele, M. Mostajabi, S. Basart, M. R. Walter, and G. Shakhnarovich, "DIODE: A dense indoor and outdoor Depth dataset," 2019, *arXiv:1908.00463*.
- [70] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "MASK R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2961–2969.
- [71] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2017, *arXiv:1611.07004*.
- [72] C. Li and M. Wand, "Precomputed real-time texture synthesis with Markovian generative adversarial networks," in *Proc. ECCV*, 2016, pp. 702–716.
- [73] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [74] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 447–456.



[75] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[76] Q. Chen, J. Xu, and V. Koltun, "Fast image processing with fully-convolutional networks," in *Proc. ICCV*, Oct. 2017, pp. 2497–2506.

[77] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in *Proc. ICLR*, 2016. [Online]. Available: <https://arxiv.org/abs/1511.07122>

[78] N. Arvanitopoulos, R. Achanta, and S. Susstrunk, "Single image reflection suppression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4498–4506.

[79] B. H. Pawan Prasad, K. S. Green Rosh, R. B. Lokesh, K. Mitra, and S. Chowdhury, "V-DESIRR: Very fast deep embedded single image reflection removal," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 2390–2399.

[80] Q. Zheng, B. Shi, J. Chen, X. Jiang, L.-Y. Duan, and A. C. Kot, "Single image reflection removal with absorption effect," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13395–13404.

[81] X. Feng, H. Ji, B. Jiang, W. Pei, F. Chen, and G. Lu, "Contrastive feature decomposition for image reflection removal," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2021, pp. 1–6.

[82] Q. Zheng, J. Chen, Z. Lu, B. Shi, X. Jiang, K.-H. Yap, L.-Y. Duan, and A. C. Kot, "What does plate glass reveal about camera calibration?" in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 3022–3032.



**ALI AMANLOU** was born in Zanjan, Iran, in October 1999. He is currently pursuing the B.Sc. degree in electrical engineering (telecommunications) with the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. His research interests include artificial intelligence, deep learning, computer vision, image processing, signal processing, and machine learning.

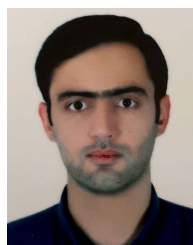


**AMIR ABOLFAZL SURATGAR** (Senior Member, IEEE) received the B.S. degree (Hons.) in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 1996, and the M.S. and Ph.D. degrees (Hons.) in control engineering from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 1999 and 2002, respectively. He is currently an Associate Professor with the Amirkabir University of Technology (Tehran Polytechnic). He has authored one

book and authored/coauthored more than 140 papers in scientific journals and conferences. He has been the Deputy Vice-President of scientific chapters in the IEEE Iran Section. He is the Founder of the Distributed and Intelligent Optimization Research Laboratory (DIORLab). His research interests include distributed optimization, computational intelligence, machine learning, machine vision, and computational neuroscience. He has been selected as one of the best Scientific-Industrial Project Manager by the Ministry of Science, Research, and Technology of Iran, in 2018.



**JAFAR TAVOOSI** was born in Ilam, Iran, in 1987. He received the B.S. degree in electrical engineering from Razi University, Kermanshah, Iran, in 2009, the M.Sc. degree in control engineering from Tabriz University, in 2011, and the Ph.D. degree in control engineering from the Amirkabir University of Technology, in 2016. He currently is an Assistant Professor with the Department of Electrical Engineering, Ilam University. He has published one book and more than 26 papers in scientific journals and conferences. His research interests include computation intelligence, fuzzy systems, and system identification.



**ARDASHIR MOHAMMADZADEH** was born in West Azerbaijan, Iran. He received the B.Sc. degree from the Sahand University of Technology, Tabriz, Iran, in July 2011, the M.Sc. degree from the KT University of Technology, Tehran, Iran, in September 2013, and the Ph.D. degree from the University of Tabriz, Tabriz, in November 2016.

In December 2017, he joined the University of Bonab, Bonab, Iran, as an Assistant Professor. He has published many papers in most reputed journals. His research interests include control theory, fuzzy logic systems, machine learning, neural networks, intelligent control, electrical vehicles, power systems control, chaotic systems, and medical systems. He is an Academic Editor of *PLOS One* and a Reviewer of several journals, such as *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, *Applied Soft Computing*, *Nonlinear Dynamics*, and many others.



**AMIR MOSAVI** (Student Member, IEEE) received the degree from London Kingston University, U.K., and the Ph.D. degree in applied informatics. He is an Alexander von Humboldt Research Fellow for big data, the IoT, and machine learning. He is a Senior Research Fellow with Oxford Brookes University. He is a Data Scientist for climate change, sustainability, and hazard prediction. He was a recipient of the Green-Talent Award, UNESCO Young Scientist Award, ERCIM

Alain Bensoussan Fellowship Award, Campus France Fellowship Award, Campus Hungary Fellowship Award, and Endeavour-Australia Leadership.

...