

Article

Transfer Learning with Convolutional Neural Networks for Rainfall Detection in Single Images

Nicla Maria Notarangelo ¹, Kohin Hirano ², Raffaele Albano ^{1,*} and Aurelia Sole ^{1,*}¹ School of Engineering, University of Basilicata, 85100 Potenza, Italy; nicla.notarangelo@gmail.com² National Research Institute for Earth Science and Disaster Resilience-NIED, Tsukuba 305-0006, Japan; hirano@bosai.go.jp

* Correspondence: raffaele.albano@unibas.it (R.A.); aurelia.sole@unibas.it (A.S.)

Abstract: Near real-time rainfall monitoring at local scale is essential for urban flood risk mitigation. Previous research on precipitation visual effects supports the idea of vision-based rain sensors, but tends to be device-specific. We aimed to use different available photographing devices to develop a dense network of low-cost sensors. Using Transfer Learning with a Convolutional Neural Network, the rainfall detection was performed on single images taken in heterogeneous conditions by static or moving cameras without adjusted parameters. The chosen images encompass unconstrained verisimilar settings of the sources: Image2Weather dataset, dash-cams in the Tokyo Metropolitan area and experiments in the NIED Large-scale Rainfall Simulator. The model reached a test accuracy of 85.28% and an F1 score of 0.86. The applicability to real-world scenarios was proven with the experimentation with a pre-existing surveillance camera in Matera (Italy), obtaining an accuracy of 85.13% and an F1 score of 0.85. This model can be easily integrated into warning systems to automatically monitor the onset and end of rain-related events, exploiting pre-existing devices with a parsimonious use of economic and computational resources. The limitation is intrinsic to the outputs (detection without measurement). Future work concerns the development of a CNN based on the proposed methodology to quantify the precipitation intensity.



Citation: Notarangelo, N.M.; Hirano, K.; Albano, R.; Sole, A. Transfer Learning with Convolutional Neural Networks for Rainfall Detection in Single Images. *Water* **2021**, *13*, 588. <https://doi.org/10.3390/w13050588>

Keywords: rain detection; transfer learning; convolutional neural networks; single image classification; rainfall; cameras

Academic Editor: Scott Curtis

Received: 23 January 2021

Accepted: 19 February 2021

Published: 24 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Floods are considered major natural disasters globally [1,2]. Furthermore, the extreme weather events stemming from climate change may increase their frequency and magnitude [3,4]. Urban and suburban landscapes are usually dominated by artificial impervious surfaces that affect the hydrological cycle [5,6]. In this context, the meteorological forcing characterization becomes essential: the identification and monitoring of hydrological state indicators, such as rainfall characteristics, are a prerequisite to determine the critical thresholds for flood triggering. These indicators contribute to shape the event and risk scenarios within the territorial context and therefore the criticality and alert levels that establish the indication or warnings to the exposed population [7,8].

For the purpose of representation of the territory for urban pluvial flood studies, precipitation measurements are required at local, fine scales with low latency to give near real-time and fine-grained information [9,10]. Urban flood risk management can be enhanced by a local scale strategic system integrating near real-time methods, innovative technologies and conventional systems [11].

Traditionally, rainfall has been assessed by employing several kinds of rain gauges [12]. Presently, most data are obtained from ground-based measurements or remote sensing: rain gauges, disdrometers, weather radars and satellites often provide limited information in terms of resolution (temporal or spatial) due to the sparseness of the ground monitoring networks that are used for both direct measurement and indirect measurement calibration.

Other problems are related to their high costs. Rain gauges are unevenly spread and frequently placed away from the urban centers. Not all observations are continuous and available to the public and often have their limitation due to the variability of the precipitation [13]. To overcome these limitations, it may be useful to employ a form of aggregation of data obtained from a dense network whose nodes are represented by numerous low-cost inaccurate sensors. Thus, several studies developed unprecedented approaches to retrieve precipitation information with low operational costs. Despite their inaccuracy, novel techniques could furnish valuable additional information in combination with conventional systems [14].

The first discussions of a camera designed to photograph rainfall emerged with the Raindrop camera [15], to support radar investigations. The idea of a Camera Based Rain Gauge was suggested by Nayar and Garg, who conducted systematic studies of the visual effects of the rain in pictures and video sequences setting the key theoretical framework [16,17] for all later research. The presence of rain produces local variations in the pixel intensity values of digital images, depending on the raindrop intrinsic characteristics, camera parameters and environment.

The literature on the perceptual properties of atmospheric precipitation and weather condition in natural images has highlighted different approaches [18–36]. The most recent advances in the field of Computer Vision (CV) and Machine Learning (ML), the new methods of analysis of digital photography and image processing, have facilitated the study of novel rain detection systems aimed at enhancing the visual aspects to increase the rainfall detection or to remove the raindrops and rain streaks from images and video frames.

The visual effects generated by the presence of rainfall have been attracting much research attention within the fields of outdoor vision. Rain-induced artifacts have typically been addressed as an element that negatively impacts the performance of systems based on video and image processing. Thus, rain detection and removal are considered as a joint pre-processing task in many CV applications. A considerable amount of literature has been published on rain removal techniques [22,28] to enhance the adverse weather degraded outdoor images or videos for different applications, such as image or video editing [33], surveillance vision system and vision-based driver assistance systems [18,24,25]. Although originally developed for visibility enhancement, the methods provide useful insights into detection techniques. Classical CV techniques were applied for the detection of dynamic weather phenomena in videos [20], whereas Deep Learning (DL) techniques—in particular Convolutional Neural Networks (CNNs)—were applied to single images [29,30]. 3D CNN systems for the stand-alone rainfall detection from video-frame sequences have proven to be effective for using existing surveillance cameras as a rain detector to decide whether to apply removal algorithms [36].

Whereas the existing literature on techniques for removing the visual effects caused by rain is fairly extensive (bearing in mind the recentness of major advances in digital photography), there is a smaller body of literature that is concerned with the field of rain measurement for hydrological purposes. A first monitoring system through pictures at the small catchment scale [21] focuses on snow precipitation visual features that differ significantly from those of the rain in terms of particles size, color and trajectory. Other opportunistic rainfall sensing methods are based classical image processing algorithms: Allamano et al. [26] estimated the rain rate values with errors of the order of $\pm 25\%$ from pictures taken at adjacent time steps; Kolte and Ghonge [31] processed video frames collected from a high definition camera; Dong et al. [34] proposed a method for real-time rainfall rate measurement from videos, counting the focused raindrops in a small depth of field to calculate a raindrop size distribution curve and estimate the corresponding rainfall rate. Jiang et al. [37] used videos acquired by surveillance cameras under specific settings and evaluated their method on synthetic numerical experiments (i.e., images processed in Adobe Photoshop) and field tests, reaching a mean absolute error of 21.8%. These techniques strongly rely on the image resolution, temporal information, acquisition device

characteristics and proper setting of its parameters in order to obtain shootings suitable for the image processing that depicts accurately detectable raindrops and rain streaks.

Adopting a ML approach, researchers have been able to devise robust algorithms for weather classification from single outdoor images, considering different weather categories: Sunny, Cloudy, Rainy [19]; Sunny and Cloudy [23,27]; Sunny, Rainy, Snowy and Haze [32]; and Sunny, Cloudy, Snowy, Rainy and Foggy [35].

In comparison with image sequences or video, single images comprise information about the recorded scene *per se* as they can stand on their own, with a consequent decrease in the amount of data that needs to be processed, stored and transferred reducing computational resources and time.

Despite the diversity of purposes, methods and epistemological fields, all the studies on the visual effects of the rain support the hypothesis that photographs can provide useful information on precipitation and that digital image acquisition devices can function as rain sensors. However, the literature focused on the use of a single image source: dashboard camera, i.e., cameras mounted under the windshield of a vehicle [18,19,24]; surveillance cameras [20,34,36,37]; outdoor scenes [23,27,30,32,33,35]; and devices with adjustable shooting setting, e.g., exposure time, F-number, depth of field, etc. [21,26,31,34,37]. Proposed methods often exploit temporal information requiring videos, frames or sequential still images [18,20,21,24,26,31,34,36,37]. Thus, the literature focuses on use cases with stringent operational requirements, and it lacks heterogeneity and generalizability and tends to be device-specific. The final goal of this research is the characterization of rainfall for creating a dense network of low-cost sensors to support the traditional data acquisition and collection methods with a relatively expeditious solution, easily embeddable into smart devices, benefiting from the aforementioned research progress and the advances in available technology, using DL techniques.

The specific aim of this work is to investigate the use of maximized number of image sources as rain detectors in order to obtain a sensing network as dense as possible providing a large informative and reliable database. As opposed to existing works, the focus of our model is the applicability on all easily accessible image acquisition devices (e.g., smartphones, general-purpose surveillance cameras, dashboard cameras, webcams, digital cameras, etc.), including the pre-existing ones. This encompasses cases where it is not possible to adjust the camera parameters to emphasize rainfall appearance or obtain shots in timelines or videos.

The proposed methodology has a number of attractive features: application simplicity; adherence to the physical-perceptual reality; cost-effective observational and computational resources. In fact, the classification task was performed on single photographs taken in extremely heterogeneous conditions by commonly used tools. The input data can be gained from different acquisition device in different lighting conditions and without special shooting settings, within the limits of visibility (i.e., human visual perception). Shots in timelines or videos are not required, so it is possible to use on pre-acquired images taken for other purposes. In an attempt to adhere physical-perceptual appearance of real rain, any form of digitally simulated rain was excluded from the input data.

In this study, we examine the case of rainfall detection, namely a supervised binary classification in single images answering the question: *Is it raining or not?*

2. Materials and Methods

The main advantage of ML method is that it uses algorithms that iteratively learn from data and allows computers to solve problems without being explicitly programmed how to do it [38]. Using this approach, researchers have been able to provide robust algorithms to solve new and complex problems, including image classification in disparate contexts.

The DL techniques, in particular CNN models, are widely used in image-classification applications so appeared especially suitable for this purpose as they are inspired by the structural and functional characteristics of the visual cortex of the animal world. A CNN-based approach was chosen since it is the state-of-art among other image recognition

methods [39]. One benefit of CNNs is that they avoid the problem of pre-processing input data since such algorithms recognize visual patterns directly in images represented by pixels. Traditional CV techniques use manually engineered feature descriptors for object detection, requiring a cumbersome process to decide which features are more descriptive or informative for the desired task. CNN models automate the process of feature engineering: the machine is trained on the given data and discovers the underlying patterns in classes of images, learning automatically the most descriptive and salient features with respect to the addressed task [40]. Deep architectures contain multiple levels of abstraction and possess the ability to learn complex, non-linear and high-dimensional functions. The general CNN structure consists of a series of layers implementing feature extraction and classification: the densely connected layers learn global patterns in their input feature space while the convolution layers learn local patterns.

The task of detecting rainfall can be achieved through the analysis of the perceptual aspects that determine whether the single photograph in heterogeneous scenes represents a condition of “presence of rain” or “absence of rain” (for obvious reasons, the two conditions are mutually exclusive classes). The scenario of interest for the proposed study is a case of supervised learning of a binary classification type [41]: the objective is to learn a mapping from known inputs x (pictures) and outputs y_i (labels), where $i \in \{1, \dots, C'\}$, with C' being the number of classes. In the binary case $C' = 2$, thus it is assumed that there are two classes corresponding to the considered weather conditions: C_1 and C_2 . The model aims to generate accurate predictions, associating a novel input (picture) to a label (C_1 or C_2 class).

The non-trivial nature of the task of interest and the small size of the available image dataset meant that one of the most effective strategies was the transfer learning [42,43] approach (Figure 1). This ML technique exploits the common visual semantics of the images [44], employing a pre-trained network for the features extraction as a base for the new classifier trained on the new dataset. A major advantage of re-purposing a network that was previously trained on a large-scale image-classification benchmark dataset such as ImageNet [45] is the portability of the learned features. Thus, the spatial hierarchy of features can work as a generic model of the visual world for novel perceptual problems [46].



Figure 1. Model setup.

2.1. Dataset Creation

The entire dataset used in the different stages of the creation of the CNN model—training, validation and test—was produced on the basis of the essential requirements inherent in the required classification task. The task of the detection of liquid precipitation was modeled as a supervised binary classification in single images. Each input instance was paired with exactly one output label describing the class. The class for dichotomization (according to the presence or absence of visible rainfall) were “With Rain” (WR) and “No Rain” (NR).

In realistic rainfall detection scenarios, the accessible images are weather degraded and exhibit a large variability of the conditions under which the images have been taken. The dataset was meant to represent unconstrained and verisimilar image settings for designing robust model capable of coping with the variations in the given images. Criteria for building the dataset were as follows: availability, weather conditions representativeness, the variety of the locations and the diversity of the capture time and lighting conditions. Each photograph depicted an outdoor scene; in the case of pictures belonging to the WR group, their visual appearance altered by the rainfall was perceptible by the human eye. The pictures with digitally synthesized rain were excluded. In particular, raindrops or streaks were not generated through processing with photo editing software (e.g., Adobe Photoshop,

GIMP, etc.) or through other computer graphics techniques, such as 3D modeling or rendering with graphics engines (e.g., Blender, Unity 3D, etc.). To increase the scene heterogeneity and have a sufficiently large and balanced number of instances, different datasets with the mentioned characteristics have been combined.

The first part of the dataset was obtained by selecting the images originally labeled as “sunny” and “rain” in the Image2Weather set [46], which is a large-scale single images dataset grouped by weather condition and freely available (Figure 2).



Figure 2. Example from Image2Weather dataset [35]: pictures labeled as rainy (left) and sunny (right).

The second source for image collection is a set of pictures taken by dashboard cameras mounted on vehicles moving around Tokyo metropolitan area from 19 August 2017 for 48 h, with an average of 4 shots captured in a minute (© NIED—National Research Institute for Earth Science and Disaster Prevention, Tsukuba, Japan). Figure 3 displays a map for the location of the images collected through the front windshield of the vehicles. The ground truth was obtained associating the rainfall rates retrieved from the high precision multi-parameter radar data [47,48] of XRAIN (eXtended RADar Information Network) operated by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan, to the images according to the capture time and GPS location. The minimum rainfall threshold used for labeling images as belonging to rainy class was 4 mm/h [49] for the veracity of the association of the radar data with the position on the ground surface and reduce the effects of smaller raindrops vertical displacement.

Finally, the third tranche of the data was built through experimental activities in the Large-scale Rainfall Simulator of the NIED located in Tsukuba (Ibaraki prefecture, Japan) shown in Figure 4. The benefit of using a large-scale rain simulator is that it allows experimental tests to be carried out in a relatively short time, reproducing events of different known intensity, including those with a rather remote occurrence frequency, such as rain showers and downpours with time high return period under controlled and repeatable conditions. This facility for hydro-geological processes simulation and measurement is the largest simulator in the world in terms of rainfall area (approximately 3000 m²) and sprinkling capacity. It can produce fairly accurate precipitation fields for rainfall intensities between 15 and 300 mm/h. The nozzles, set up at a height of 16 m above the ground surface, are capable of reproducing the natural rainfall constant speed and generating raindrops with a diameter ranging from 0.1 to 6 mm.

During the experiments, the produced intensity was set to a constant value for 15 min intervals; the nominal values for each interval ranged from 20 to 150 mm/h and were assumed as ground truth. For the purpose of the heterogeneity of the scenes, the photographs (Figure 4) were shot with 5 different devices: Canon XC10 (Oita Canon Inc., Oita, Japan), Sony DSC-RX10M3 (Sony Corporation, Tokyo, Japan), Olympus TG-2 (Olympus Corporation, Tokyo, Japan), XiaoYI YDXJ 2 (Xiaoyi Technology Co. Ltd., Shanghai, China) and the smartphone XiaoMi MI8 (Xiaomi Communications Co., Ltd., Beijing, China).

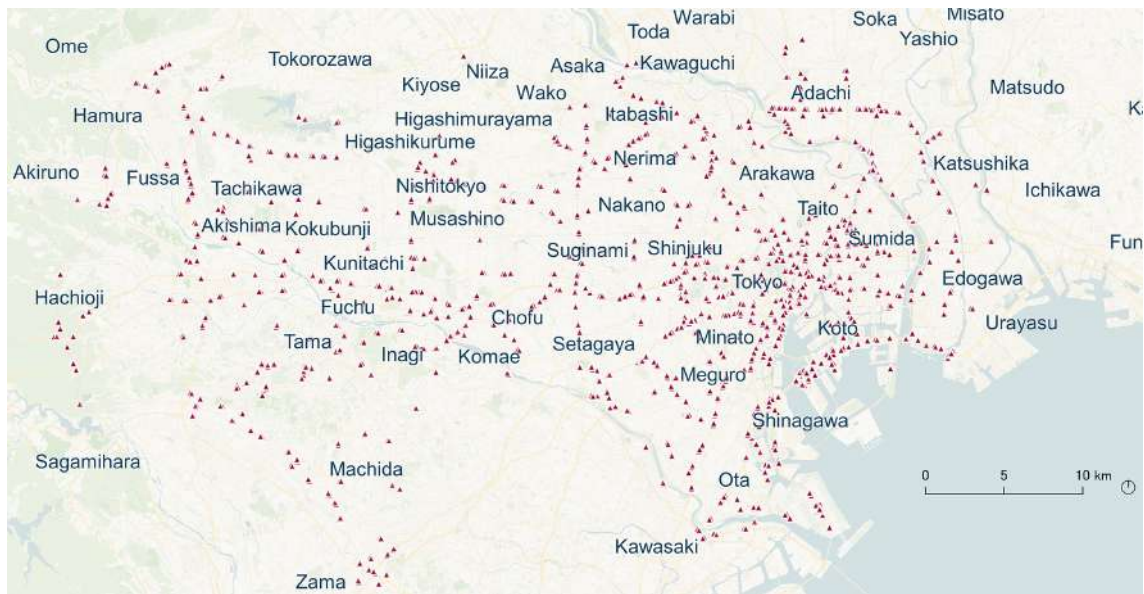


Figure 3. Geolocation of the images taken from the dashboard cameras.

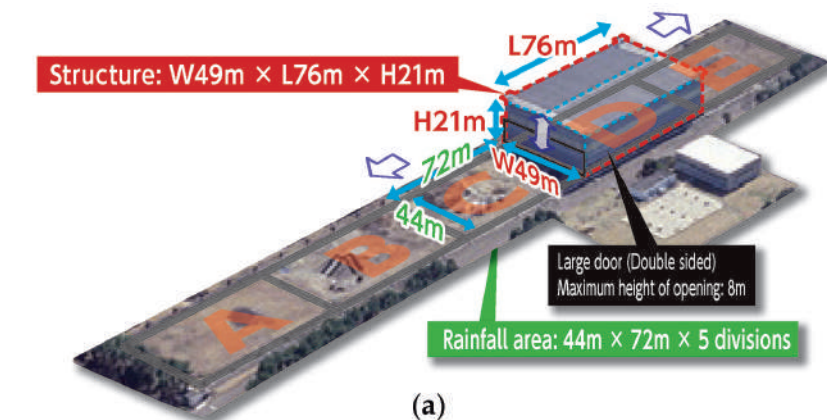


Figure 4. (a) NIED Large-scale Rainfall Simulator located in Tsukuba (Ibaraki prefecture, Japan) ©NIED. (b) Example of the dataset built during the NIED Large-scale Rainfall Simulator experiments: pictures labeled as With Rain (left) and No Rain (right).

The total dataset consisted of 7968 color images coming from the different devices saved in the JPEG format (pixel resolution from 333×500 to 5472×3648 pixels). All the collected images were labeled and organized into the necessary non-overlapping categories [50], as shown in Table 1. The number of instances of the two classes was properly balanced.

Table 1. Dataset creation: label division and number of labeled images per dataset.

Label	Dataset	# of Pictures			
		(Total)	Train	Validation	Test
WR (presence of rainfall)	Image2weather	1369	823	273	273
	NIED car pictures dataset	1870	1123	374	374
	NIED Rainfall Simulator Experiment	744	448	148	148
	TIM dataset (Experiment)	1180	-	-	1180
NR (absence of rainfall)	Image2weather	1369	823	273	273
	NIED car pictures dataset	1870	1123	374	374
	NIED Rainfall Simulator Experiment	744	448	148	148
	TIM dataset (Experiment)	1180	-	-	1180

To mitigate overfitting caused by the small number of images, data augmentation was adopted as a strategy to artificially create new training examples from the existing ones [51]. The random geometric transformations were chosen so as to be isometric and guarantee the physical compatibility with the natural appearance of meteorological phenomena. Other data augmentation methods (e.g., color space transformations, random erasing, mixing images, etc.) were avoided since they compromise the visual effects of the rain in pictures. The new pictures were generated by a mirror-reversal of an original across a vertical axis (horizontal flip or reflection) and/or a 2 degrees angle rotation.

2.2. Model Architecture

The deep learning model was entirely implemented with open source tools, namely R programming language with the Integrated Development Environment RStudio using Keras framework and Tensorflow engine backend [46].

A transfer learning approach was adopted for the features extraction, choosing the VGG16 network [52] as convolutional base because of its characteristics of generality and portability of the learned features, excluding the densely connected classifier on its top. The VGG16 network was trained on 1.4 million labeled images of the ImageNet dataset and 1000 different classes (mainly animals and everyday objects) achieving a classification accuracy of 92.7%. Since the convolutional base of VGG16 has 14,714,688 parameters, it was frozen before the compiling and training of the new network. Freezing was necessary to prevent the weights from being updated during the training of the new network so as to preserve the representations that were previously learned and avoid overfitting that can result from fine-tuning [43]. The model was extended by adding new layers on top (Table 2), in order to obtain a new classifier capable of generating predictions with the two desired output classes (predict whether the picture represents rainy conditions, WR, or not, NR). The dropout layer, which randomly drops some input units and their connections with probability $1 - p$, was used as a regularization technique to reduce possible data overfitting problems in the learning phase so as to preserve the algorithm generalizability.

Table 2. Model architecture.

Layer (Type)	Output Shape	Param. #
vgg16 (Model)	(None, 6, 6, 512)	14,714,688
flatten (Flatten)	(None, 18,432)	0
dense (Dense)	(None, 256)	4,718,848
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257
Total params: 19,433,793 (14,714,688 frozen)		

3. Results

3.1. Training and Validation

The labeled pictures of the training set (Table 1) constituted known examples of ordered pairs of input (image) and output (WR or NR class) to feed the network so that the final model can make predictions starting from inputs.

The CNN was initialized with random values of the trainable parameters. Then, the optimizer updated automatically—for each elaborated instance—the values of the weights implementing the error backpropagation algorithm. The optimization procedure modified the training parameters of the model in order to maximize the prediction accuracy and minimize the cost function. The dropout rate was set to 25, meaning one in four inputs was randomly excluded from each update cycle during the training phase. The loss of binary cross entropy between the training data and the model's predictions was used as the cost function [44]. A disjoint validation set provided an unbiased evaluation of the model fit on the training dataset while tuning the model's hyperparameters, thus preserving its generalization ability.

The training configuration setup can be summarized as follows:

- Data Augmentation: horizontal flip, 2° rotation
- Loss: binary cross-entropy as cost function for monitoring and improving the model skills
- Optimizer: *RMSprop Optimizer* that implements the RootMean Square Propagation (RMSprop) algorithm [53] to update the neural network
- Metric: accuracy. The amount of correctly classified instances
- Learning rate: $lr = 1 \times 10^{-5}$

Figure 5 illustrates the predictive effectiveness of the network plotting the trends of the accuracy (*acc*) and loss (*loss*) values over the subsequent iterations (*epoch*) calculated on the known data, i.e. training, and *unknown* ones, i.e. validation. During the first 15 epochs, accuracy increased while the loss decreased, constantly improving without abrupt changes the training stability. After 17 epochs, the accuracy reached a stable value of $\approx 87\%$ for the training set and $\approx 84\%$ for the hold-out validation set. The loss stalled at ≈ 0.30 for the training set and ≈ 0.33 for the validation set. For the subsequent epochs, the plots indicate that the training and validation curves stopped improving significantly. The small gap between the curves relative to the two subsets indicates that the network has converged reaching a point of stability without over-learning of the training data.

The training was stopped after 30 epochs, to avoid the degradation of the performance of the model [44] due to possible overfitting which is graphically identifiable as a divergence between the curves of training and validation for each metric.

During the 30 epochs—complete presentations of the entire training and validation datasets—the learning rate was set to a fairly low value (1×10^{-5}) to maintain stability [44]. The trained model reached an accuracy of 88.95% on the training set and 85.47% on the validation set and a loss (Murphy 2012) of 0.27 and 0.33, respectively.

The Loss (Binary Cross Entropy, i.e. CE) was calculated as follows:

$$CE = - \sum_{i=1}^{C'=2} t_i \log(f(s_i)) = -t_1 \log(f(s_1)) - (1 - t_1) \log(1 - f(s_1)) \quad (1)$$

where t_i and s_i are, respectively, the ground-truth and the raw score for each class i in C and

$$f(s_i) = \frac{1}{1 + e^{-s_i}} \quad (2)$$

is the logistic sigmoid activation function which is necessary for the output to be interpreted as a probability.

The case of interest is a binary classification problem, where $C' = 2$. It is assumed that there are two classes, C_1 and C_2 ; thus, $t_1 \in [0, 1]$ and $f(s_1) \in [0, 1]$ are the ground truth and

the probability value for C_1 , and $t_2 = 1 - t_1$ and $f(s_2) = 1 - f(s_1)$ are the ground truth and the output score for C_2 , as the classes are mutually exclusive.

The values are significantly far from a random prediction in a balanced binary problem: modeling the random predictor as a “fair coin” [41,54], where the tossing is a discrete-time stochastic process giving a binary output, the non-informative values correspond to accuracy of 50% and loss of 0.69.

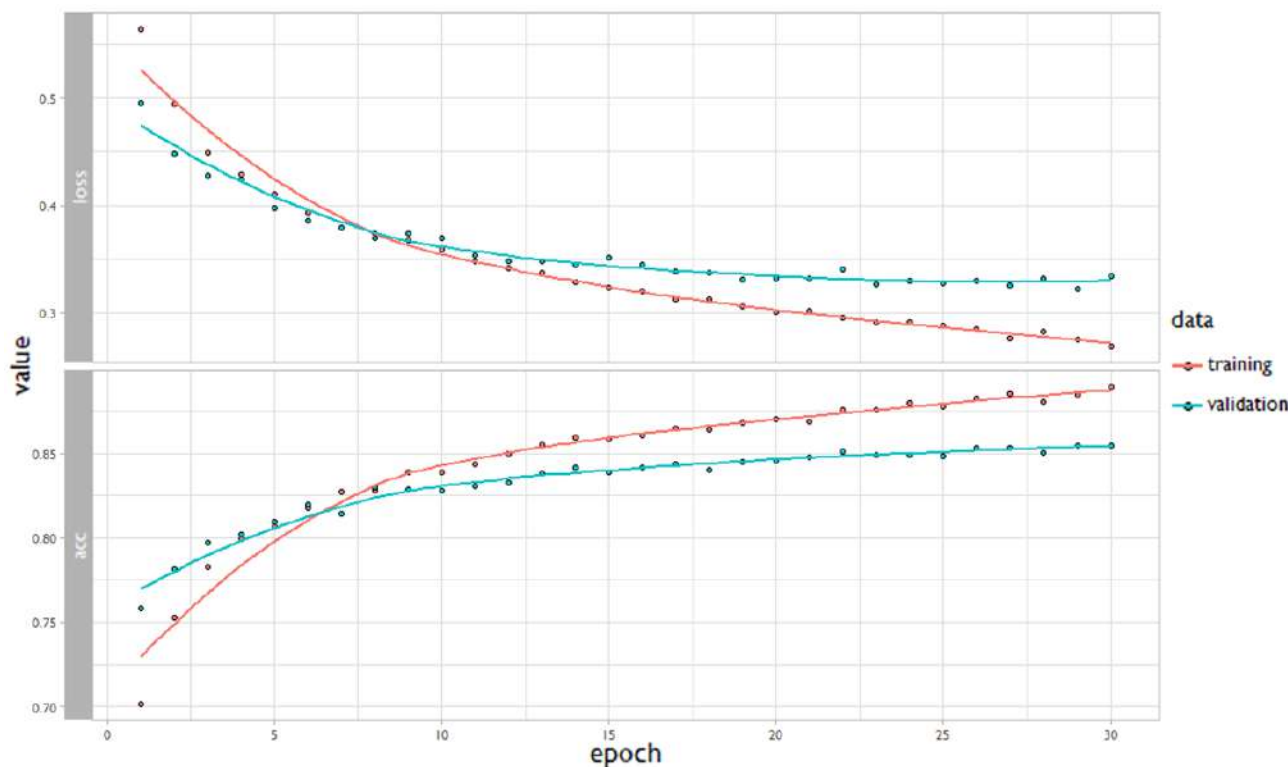


Figure 5. Training and validation: overall accuracy and loss.

3.2. Test

To evaluate the performance of the proposed algorithm, the trained model was applied to the test dataset to give an unbiased estimate of model skill. The test set is independent of the training and validation datasets, but follows the same probability distribution. To assess the quality of the predictions on new pictures, the confusion matrix (Table 3) was constructed over the entire test set, opposing instances in a predicted class (model response) against instances in an actual class.

Given the WR label (presence of rain) as the positive class and NR (absence of rain) as the negative class, the numbers of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) were counted to calculate useful metrics.

Table 3. Confusion matrix for the test.

		Reference	
		Positive = WR	Negative = NR
Prediction	Positive = WR	719 of TP	158 of FP
	Negative = NR	76 of FN	637 of TN

The accuracy, with a significant p -value (equal to 2×10^{-16}), and the loss of entropy calculated on the test set are, respectively, 85.28% and 0.34. The accuracy and loss are significantly different from a random prediction and consistent with the metrics obtained

in the training and evaluation phase. To measure the rain detection performance, several metrics are used and compared to reference values [41,55,56], as reported in Table 4.

From the results, it is evident that the model exhibits good overall predictive reliability. The sensitivity, >90%, is higher than the specificity, >80%. That means the test is responsive to rainy condition detection but prone to the risk of reporting false positives. In the application field of interest, i.e., urban flood risk management, high sensitivity is desirable: missing cases of rainfall presence could lead to delays in early warning systems. The detection will easily mark rainy pictures, ensuring that dangerous pluviometric forcing is recognized. On the other hand, the classification was performed on balanced class, although in reality a condition of absence of rain is more likely to occur. Hence, specificity is important to avoid unnecessary alerts.

The false positive occurrence may be explained by the presence of lens flare caused by sun-rays and artificial lighting which produces artifacts that may resemble raindrops accumulated on the lens. Closer examination of the results reveals some ambiguity in the visual appearance of the misclassified pictures and hence a limitation of our system: it does not suit scenes that are also misleading for human vision.

Table 4. Model evaluation metrics.

Metric		Value Test Set	Reference Values
Overall accuracy	$\frac{(TP + TN)}{(TP + FP + FN + TN)}$	85.28%	worst = 0% best = 100%
Cross Entropy Loss	$-\sum_{i=1}^{C'-2} t_i \log(f(s_i))$	0.3400635	perfect ≈ 0 good < 0.69 bad ≥ 0.69
Sensitivity–Recall	$\frac{TP}{TP + FN}$	90.44%	worst = 0% best = 100%
Specificity	$\frac{TN}{FP + TN}$	80.13%	worst = 0% best = 100%
Precision	$\frac{TP}{TP + FP}$	81.98%	worst = 0% best = 100%
F ₁	$\frac{(1 + \beta^2) \times precision \times recall}{(\beta^2 \times precision) + recall}$ $= \frac{2 \times precision \times recall}{precision + recall}$	0.8600	worst = 0 best = 1
Matthews correlation coefficient MCC	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$	0.7094	worst = -1 best = +1

3.3. Model Deployment

To estimate the skills of the proposed model in a real-world operational setting, a completely new set of input images was built. The first test set was obtained from the same sources of the training and validation sets (Table 1), whereas, for the model deployment, the source of the picture was a Reolink surveillance camera installed by TIM s.p.a. within the “Bari—Matera 5G project” (<http://www.barimatera5g.it/>, accessed on 1 January 2021) without any special setting aimed to enhance the precipitation visibility. In fact, the webcam was formerly set for the Public Safety case using 5G network connectivity testing. The camera frames piazza Vittorio Veneto—the main square of Matera (Italy)—and its hypogeum (Figure 6).



Figure 6. Geolocation of the TIM surveillance camera in Matera, Italy.

The experiment test set is strongly independent of the training and validation datasets. It contains 2360 unseen examples, 1180 instances for the WR class and 1180 for the NR class. The resolution of the color picture was 2560 X 1920pixels (Table 1). The ground truth was created by manually labeling the pictures that displayed precipitation streaks or raindrops and selecting the same number of frames that displayed a clean background, with as many lighting conditions as possible (i.e., with diverse capture time). The presence of rain was verified with precipitation data collected both from the JAXA Global Rainfall Watch—GSMaP Project (<https://sharaku.eorc.jaxa.jp/GSMaP>, accessed on 1 January 2021) and the Centro Funzionale Decentrato Basilicata (<http://www.centrofunzionalebasilicata.it/>, accessed on 1 January 2021).

The learned model was used to generate predictions on the test TIM dataset. Figure 7 shows the Receiver Operating Characteristic (ROC) curve and Youden index [57,58] calculated on the output scores. The dashed grey line represents the non-informative values (chance line) of a random classifier. The ROC curve (blue line) plotted the dependency of true positive rate (Sensitivity) on false positive rate (it equal 1—Specificity) obtained at various thresholds. The area under the ROC curve is 0.911 with confidence interval 95% = (0.899, 0.923), suggesting that the model predicts rainfall presence very well. Using the Youden index point, the optimal cutoff value was set to 0.6.

Both TIM image set and binary test set share a significant outcome.

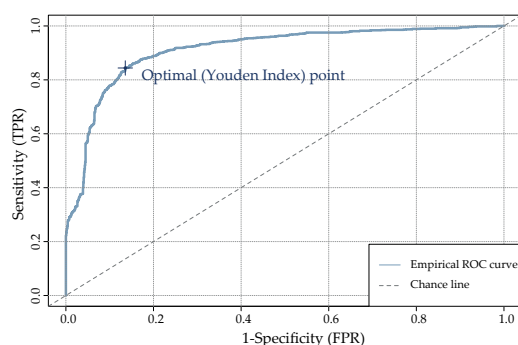


Figure 7. The ROC curve for prediction model in the TIM surveillance camera sample.

The confusion matrix is given in Table 5.

Table 5. Confusion matrix for the TIM dataset.

		Reference	
Prediction	Positive = WR	Positive = WR 981 of TP	Negative = NR 199 of FP
	Negative = NR	152 of FN	1028 of TN

The overall accuracy, with a p-value equal to 4×10^{-14} , and the loss of entropy computed on the TIM image set are 85.13% and 0.396. The metrics values were similar (Table 6), exhibiting a rather high predictive effectiveness [41,55,56].

Table 6. Model evaluation metrics: test and deployment.

Metric	Value Test Set	Value TIM Set
Overall accuracy	85.28%	85.13%
Cross Entropy Loss	0.3400635	0.3960878
Sensitivity—Recall	90.44%	83.14%
Specificity	80.13 %	87.12%
Precision	81.98%	86.58%
F ₁	0.8600	0.8482
Matthews correlation coefficient MCC	0.7094	0.7031

Compared to the test results, the model deployed on the TIM camera pictures discloses lower sensitivity, $\approx 83\%$, but higher specificity, $\approx 87\%$, and precision, $\approx 86\%$. That means that the pictures marked as relevant instances (WR) were actually relevant, so the alerts are likely to be true. On the other hand, the higher false negative rate may underestimate the pluviometric forcing.

Considering an optimal balance of recall and precision, F1 score is close to the ideal value of 1 in both the test evaluation and the experiment evaluation. The model withstood the drastic change in the datasets, as the images were completely new.

Figure 8 shows a frame taken by TIM surveillance camera in rainy condition and a visual explanation of the CNN classification, the Gradient-weighted Class Activation Mapping (Grad-CAM) [59]. The heatmap highlights the specific discriminative regions within the input image that are relevant for the classifier predictions. It is interesting to note that the parts of the image judged more *rain-like* (strongly activated in the Grad-CAM) by our CNN are the portions of the scene obstructed by the raindrops accumulated on the lens of the surveillance camera.



Figure 8. Example of a frame in rainy condition (left) and its grad-CAM (right).

4. Discussion

The results of the model test and the deployment on the real use case are encouraging for the binary classification problem that detects the presence or the absence of precipitation.

To assess the classification effectiveness of our study, we compare our method with other state-of-the-art methods for rain detection from a single image [19,32,35]. Because there is no publicly available implementation of these algorithms, the comparison was made in terms of the performance metrics reported in the literature. The model and features used, image source, labels for classification, overall accuracy and rain class sensitivity are given in Table 7.

Table 7. Comparison between results of the current study and results from other studies.

Method	Model and Features Used	Image Source	Classes	Overall Accuracy	Sensitivity (Rain)
Yan et al. [19]	Real AdaBoost with histogram of gradient amplitude, HSV color histogram, road information	In-vehicle vision system	Sunny, Cloudy, Rainy	91.92%	90.41%
		Crawled outdoor scenes, as reported in [32]		18.89%	-
Zhang et al. [32]	Multiple kernel learning with engineered features (sky, shadow, rain streak, snowflake, dark channel, contrast, saturation)	Crawled outdoor scenes	Sunny, Rainy, Snowy, Haze	71.39%	67%
Chu et al. [35]	Random Forest with time, RGB color histogram, Gabor wavelet, intensity histogram, local binary pattern, cloud, haze, and contrast features	Crawled outdoor scenes	Sunny, Cloudy, Snowy, Rainy, Foggy.	76.8%	68%
Proposed method	CNN with pre-learned filters	Dash-cams, consumer cameras, outdoor scenes from [35], smartphone	With Rain, No Rain.	85.28%	90.44%
		Surveillance camera		85.13%	83.14%

Table 7 shows that the AdaBoost-based model [19] achieves the best accuracy over all other methods. By comparing sensitivity obtained by our model on the test set with that in [19], we see similar values. On the other hand, the model in [19] uses engineered features based on in-vehicle vision system (i.e., *road information*) for classification. As a consequence, the model may experience abrupt degradation of accuracy for unconstrained scenarios, as shown in the results of generic outdoor scenes classification reported in [32]. The results on less restrictive settings reveal our new methods offers performance advantages in both test and deployment real-world scenarios; the values of accuracy and sensitivity obtained by our model in both cases are higher than the values obtained with the approaches by Yan et al. [19], Zhang et al. [32], Chu et al. [35]. This result may be explained by the two main differences between ours and the aforementioned studies. First, our method uses pre-learned filters for feature extraction (frozen VGG16 parameters, Table 2), so it does not require manually made feature descriptors for object detection, which are susceptible to recall bias. Secondly, the trainable part of our network (Table 2) learns from a dataset that exhibits large variability and includes both static and in-vehicle cameras, whereas all the studies tend to be device-specific.

Hence, the results achieved surpass the earlier work in this area in terms of input heterogeneity, as the model can handle images captured with both static or moving cameras without significant accuracy loss. Our algorithm is simple, flexible, robust and leaves no requirement for any other step.

The proposed approach allows using heterogeneous and easily producible input data: the images can be acquired from disparate sources in different lighting conditions, without particular requirements in terms of the shooting settings (although some parameters may favor the visibility of the precipitation), and, since each image is classified independently, the detection does not require any sequence or continuous shooting or videos. Once the training phase is over, computational times are extremely reduced.

This approach is especially useful in urban areas where video and image capture devices are ubiquitous and measurement methods such as rain gauges encounter installation difficulties or operational limitations or in contexts where there is no availability of remote sensing data.

The result is a fairly rapid, simple and expeditious solution. As demonstrated in the experiment on a real-world use case, it is immediately applicable at an operational level and easily embeddable in intelligent devices. The experimental activity within the “Bari Matera 5G” project provides a promising perspective on the Smart Environmental Monitoring using 5G connectivity. The system can be managed locally, remotely or via cloud by programming electronic devices (such as smart cameras or IoT devices) for recording and transmitting the rain data. The model can retrieve rainfall information by processing data collected from different sources located in a fixed place (surveillance cameras) or moving (smartphones, dash-cams and action-cams), contributing to the traditional monitoring networks and forecast systems.

These findings corroborate the idea in [27] that CNN-based classification outperforms engineered learning representation for weather-related vision tasks due to the CNN model’s ability to finding the nonlinear mapping and taking into account extensive factors including low-level abstraction using pre-trained convolutional base (transfer learning).

The major limitation of the proposed approach is intrinsic to the output results: in fact, the detection of the presence of precipitation (without a quantification) is a necessary condition but not sufficient for the creation of a rain gauge network based on cameras. An additional uncontrolled factor, common to all vision based methods, is the possibility to have ambiguity in the pictures: artifacts due to dirty lenses and flare can lead to false positives or the visual effects of precipitation may also often be inconspicuous for human observers. Misclassification error can be compensated for by many observations. The straightforwardness of our approach will make it easy and virtually cost-free to embed the proposed CNN in existing outdoor vision systems to increase the spatial density of rain observations. The rainfall detection is performed on single images that can be collected in a very short time; therefore, very high temporal resolution is possible in principle.

The proposed transfer learning technique can also be combined with incremental learning algorithms [60,61]. Incremental learning extends the existing model’s knowledge without access to the original data, by continuously learning from new input data while preserving the previously acquired knowledge. Due to this, the network has the potential to enhance the robustness to rainfall intensity (e.g., light rainfall rates), background and other environmental effects, as well as accommodate different use case scenarios (e.g., setting site-specific rainfall thresholds, detecting snow precipitation).

5. Conclusions

In this study, a DL method for detecting rainfall from single photographs taken in very generalizable conditions with no need for parameter adjustment, photometric constraint or image pre-processing was proposed. It can be applied to any outdoor scene within the limits of the human visibility.

Taken together, the results suggest that using a DL method is an efficient strategy for the CV task of gathering rainfall information using generic cameras as low-cost sensors. This finding, while preliminary, can help define novel strategies for meteorological-hydrological studies on a local scale in urban areas. Despite its exploratory nature, the proposed approach offers a first near real-time tool based on artificial vision techniques which can contribute to creating a support sensors network for the characterization of rainfall by monitoring the onset and end of rain-related events, with a parsimonious approach in terms of economic and computational resources.

The limited operational requirements allowed the proposed model to be immediately applicable in experiments on real use cases. Our CNN can easily exploit pre-existing devices and is practical for systems that need to be implemented quickly. The experimentation raises the possibility that the use of real rain images captured by already installed surveil-

lance cameras can help in setting up a monitoring and warning system of the pluviometric forcing triggering rain-related events.

This information can be used to develop targeted interventions aimed at building-up an information infrastructure integrating low-cost sensor and crowdsourced data that suits the future needs in terms of spatial and temporal resolution, scalability, heterogeneity, complexity and dynamicity.

Future work concerns the development and experimentation of another CNN network based on the proposed methodology to allow the quantitative characterization of the precipitation. The rainfall intensity will be classified into ranges; the dataset will be expanded by the number of instances and improved according to criteria of non-ambiguity and availability of known rainfall intensity values; and the model will be properly re-adapted, evaluated, calibrated and fine-tuned to meet the needs of hydro-meteorological monitoring.

Author Contributions: Conceptualization, R.A. and A.S.; methodology, N.M.N.; software, N.M.N.; validation, N.M.N. and R.A.; formal analysis, N.M.N.; investigation, N.M.N. and K.H.; resources, K.H.; data curation, N.M.N. and K.H.; writing—original draft preparation, N.M.N.; writing—review and editing, N.M.N., R.A., K.H. and A.S.; visualization, N.M.N.; supervision, A.S.; project administration, A.S.; and funding acquisition, A.S. and R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Regione Basilicata, within the Engineering School of University of Basilicata PhD Program “Engineering for Innovation and Sustainable Development”-Industry 4.0, topic: Enhancing flood risk management, through near real-time methods and technologies, for risk communication at local scale.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from National Research Institute for Earth Science and Disaster Resilience (NIED). Restrictions apply to the availability of these data, which were used under license for this study.

Acknowledgments: The assistance provided by the staff of the National Research Institute for Earth Science and Disaster Resilience (NIED), the Storm, Flood and Landslide Research Division and the Large-scale Rainfall Simulator with the different datasets and the experiments was greatly appreciated. The authors gratefully acknowledge TIM s.p.a. for providing the Matera camera frames for this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CV	Computer Vision
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
NR	No Rain
WR	With Rain
ROC	Receiver Operating Characteristic

References

1. EM-DAT. *EM-DAT | OFDA/CRED International Disaster Database*; EM-DAT: Brussels, Belgium, 2019.
2. Pesaresi, M.; Ehrlich, D.; Kemper, T.; Siragusa, A.; Florczyk, A.J.; Freire, S.; Corbane, C. *Atlas of the Human Planet 2017: Global Exposure to Natural Hazards*; European Commission, Joint Research Centre: Luxembourg, 2017.

3. de Moel, H.; van Alphen, J.; Aerts, J.C.J.H. Flood maps in Europe—Methods, availability and use. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 289–301. [[CrossRef](#)]
4. IPCC. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2012.
5. Fletcher, T.D.; Andrieu, H.; Hamel, P. Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art. *Adv. Water Resour.* **2013**, *51*, 261–279. [[CrossRef](#)]
6. Griffiths, J.A.; Singh, S.K. Urban Hydrology in a Changing World. In *Hydrology in a Changing World: Challenges in Modeling*; Singh, S.K., Dhanya, C., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 73–88. [[CrossRef](#)]
7. Scarpino, S.; Albano, R.; Cantisani, A.; Mancusi, L.; Sole, A.; Milillo, G. Multitemporal SAR Data and 2D Hydrodynamic Model Flood Scenario Dynamics Assessment. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 105. [[CrossRef](#)]
8. Sole, A.; Giosa, L.; Albano, R.; Cantisani, A. The Laser scan data as a key element in the hydraulic flood modelling in urban areas. *ISPRS Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2013**, *XL-4/W1*, 65–70. [[CrossRef](#)]
9. Einfalt, T.; Arnbjerg-Nielsen, K.; Spies, S. An enquiry into rainfall data measurement and processing for model use in urban hydrology. *Water Sci. Technol.* **2002**, *45*, 147–152. [[CrossRef](#)]
10. Schilling, W. Rainfall data for urban hydrology: What do we need? *Atmos. Res.* **1991**, *27*, 5–21. [[CrossRef](#)]
11. Albano, R.; Sole, A. Geospatial Methods and Tools for Natural Risk Management and Communications. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 470. [[CrossRef](#)]
12. Strangeways, I. A history of rain gauges. *Weather* **2010**, *65*, 133–138. [[CrossRef](#)]
13. Kidd, C.; Becker, A.; Huffman, G.J.; Muller, C.L.; Joe, P.; Skofronick-Jackson, G.; Kirschbaum, D.B. So, How Much of the Earth’s Surface Is Covered by Rain Gauges? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 69–78. [[CrossRef](#)]
14. Tauro, F.; Selker, J.; Giesen, N.v.d.; Abrate, T.; Uijlenhoet, R.; Porfiri, M.; Manfreda, S.; Caylor, K.; Moramarco, T.; Benveniste, J.; et al. Measurements and Observations in the XXI century (MOXXI): Innovation and multi-disciplinarity to sense the hydrological cycle. *Hydrol. Sci. J.* **2018**, *63*, 169–196. [[CrossRef](#)]
15. Jones, D.; Dean, L.A. *A Raindrop Camera*; Illinois State Water Survey: Champaign, IL, USA, 1953.
16. Garg, K.; Nayar, S.K. Vision and Rain. *Int. J. Comput. Vis.* **2007**, *75*, 3–27. [[CrossRef](#)]
17. Nayar, S.K.; Garg, K. When Does a Camera See Rain? *Comput. Vis. Int. Conf.* **2005**, *2*, 1067–1074. [[CrossRef](#)]
18. Roser, M.; Geiger, A. Video-based raindrop detection for improved image registration. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, Japan, 27 September–4 October 2009; pp. 570–577.
19. Yan, X.; Luo, Y.; Zheng, X. Weather Recognition Based on Images Captured by Vision System in Vehicle. In *Advances in Neural Networks*; Yu, W., He, H., Zhang, N., Eds.; Springer: Berlin, Germany, 2009; pp. 390–398.
20. Bossu, J.; Hautière, N.; Tarel, J.P. Rain or Snow Detection in Image Sequences Through Use of a Histogram of Orientation of Streaks. *Int. J. Comput. Vis.* **2011**, *93*, 348–367. [[CrossRef](#)]
21. Parajka, J.; Haas, P.; Kirnbauer, R.; Jansa, J.; Blöschl, G. Potential of time-lapse photography of snow for hydrological purposes at the small catchment scale. *Hydrol. Process.* **2011**, *26*, 3327–3337. [[CrossRef](#)]
22. Tripathi, A.K.; Mukhopadhyay, S. Removal of rain from videos: A review. *Signal Image Video Process.* **2014**, *8*, 1421–1430. [[CrossRef](#)]
23. Lu, C.; Lin, D.; Jia, J.; Tang, C. Two-Class Weather Classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *39*, 3718–3725. [[CrossRef](#)]
24. Nashashibi, F.; De Charette De La Contrie, R.; Lia, A. Detection of Unfocused Raindrops on a Windscreen using Low Level Image Processing. In Proceedings of the International Conference on Control, Automation, Robotics and Vision, Singapore, 7–10 December 2010; p. P0655.
25. Hassim, R.; Bade, A. Taxonomy of rain detection and rain removal techniques. *Trans. Sci. Technol.* **2015**, *2*, 28–35.
26. Allamano, P.; Croci, A.; Laio, F. Toward the camera rain gauge. *Water Resour. Res.* **2015**, *51*, 1744–1757. [[CrossRef](#)]
27. Elhoseiny, M.; Huang, S.; Elgammal, A. Weather classification with deep convolutional neural networks. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 3349–3353.
28. Shorman, S.; Ali Pitchay, S. A Review of Rain Streaks Detection and Removal Techniques for Outdoor Single Image. *J. Eng. Appl. Sci.* **2016**, *11*, 6303–6308.
29. Fu, X.; Huang, J.; Ding, X.; Liao, Y.; Paisley, J. Clearing the Skies: A deep network architecture for single-image rain removal. *IEEE Trans. Image Process.* **2017**, *26*, 2944–2956. [[CrossRef](#)]
30. Yang, W.; Tan, R.T.; Feng, J.; Liu, J.; Guo, Z.; Yan, S. Joint Rain Detection and Removal via Iterative Region Dependent Multi-Task Learning. *arXiv* **2016**, arXiv:1609.07769.
31. Kolte, G.; Ghonge, a.P.A. An Image Processing based Raindrop Parameter Estimation. *Int. J. Eng. Trends Technol.* **2016**, *31*, 73–77. [[CrossRef](#)]
32. Zhang, Z.; Ma, H.; Fu, H.; Zhang, C. Scene-free multi-class weather classification on single images. *Neurocomputing* **2016**, *207*, 365–373. [[CrossRef](#)]
33. Katre, R.; Dodkey, N. Rain Streaks Detection and Removal in Image based on Entropy Maximization and Background Estimation. *Int. J. Comput. Appl.* **2017**, *164*, 17–20. [[CrossRef](#)]

34. Dong, R.; Liao, J.; Li, B.; Zhou, H.; Crookes, D. Measurements of Rainfall Rates From Videos. In Proceedings of the 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017), Shanghai, China, 14–16 October 2017.
35. Chu, W.T.; Zheng, X.Y.; Ding, D.S. Camera as weather sensor: Estimating weather information from single images. *J. Vis. Commun. Image Represent.* **2017**, *46*, 233–249. [[CrossRef](#)]
36. Haurum, J.B.; Bahnsen, C.H.; Moeslund, T.B. Is it Raining Outside? Detection of Rainfall using General-Purpose Surveillance Cameras. *arXiv* **2019**, arXiv:1908.04034.
37. Jiang, S.; Babovic, V.; Zheng, Y.; Xiong, J. Advancing Opportunistic Sensing in Hydrology: A Novel Approach to Measuring Rainfall With Ordinary Surveillance Cameras. *Water Resour. Res.* **2019**, *55*. [[CrossRef](#)]
38. Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* **1959**, *3*, 210–229. [[CrossRef](#)]
39. Srinivas, S.; Sarvadevabhatla, R.K.; Mopuri, K.R.; Prabhu, N.; Kruthiventi, S.S.S.; Babu, R.V. A Taxonomy of Deep Convolutional Neural Nets for Computer Vision. *Front. Robot. AI* **2016**, *2*. [[CrossRef](#)]
40. Mahony, N.O.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Velasco-Hernandez, G.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. Traditional Computer Vision. *arXiv* **2020**, arXiv:1910.13796. [[CrossRef](#)]
41. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; Adaptive Computation and Machine Learning Series; MIT Press: Cambridge, MA, USA, 2012.
42. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
43. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the 27th International Conference on Neural Information Processing Systems—Volume 2 (ICIP), Montreal, QC, Canada, 8–13 December 2015; pp. 3320–3328. [[CrossRef](#)]
44. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
45. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
46. Chollet, F.; Allaire, J.J. *Deep Learning with R*, 1st ed.; Manning Publications Co.: Greenwich, CT, USA, 2018.
47. Hirano, K.; Maki, M.; Maesaka, T.; Misumi, R.; Iwanami, K.; Tsuchiya, S. Composite rainfall map from C-band conventional and X-band dual-polarimetric radars for the whole of Japan. In Proceedings of the 8th European Conference on Radar in Meteorology and Hydrology, Garmisch-Partenkirchen, Germany, 1–5 September 2014.
48. Hirano, K.; Maki, M. Imminent Nowcasting for Severe Rainfall Using Vertically Integrated Liquid Water Content Derived from X-Band Polarimetric Radar. *J. Meteorol. Soc. Jpn.* **2018**, *96A*, 201–220. [[CrossRef](#)]
49. Suzuki, A.; Yamada, T.J. Rainfall re-estimation associated with differences in the detection time of rainfall between X-band mp radar and raingauge. *J. Jpn. Soc. Civ. Eng. Ser.* **2015**, *71*, I257–I262. (In Japanese) [[CrossRef](#)]
50. Russell, S.J.; Norvig, P.; Davis, E. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall Series in Artificial Intelligence; Prentice Hall: Upper Saddle River, NJ, USA, 2010.
51. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
52. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556v6.
53. Tieleman, T.; Hinton, G. Lecture 6.5 rmsprop: Divide the gradient by a running average of its recent magnitude. *Neural Netw. Mach. Learn.* **2012**, *4*, 26–31.
54. Kerrich, J.E. An Experimental Introduction to the Theory of probability. *Nature* **1946**, *158*, 360. [[CrossRef](#)]
55. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)]
56. Zheng, A. *Evaluating Machine Learning Models*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2015.
57. Powers, D.M.W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**, arXiv:2010.16061.
58. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [[CrossRef](#)]
59. Selvaraju, R.R.; Das, A.; Vedantam, R.; Cogswell, M.; Parikh, D.; Batra, D. Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization. *arXiv* **2016**, arXiv:1610.02391. [[CrossRef](#)]
60. Ade, R.R.; Deshmukh, P.R. Methods for Incremental Learning: A Survey. *Int. J. Data Min. Knowl. Manag. Process.* **2013**, *3*, 119–125. [[CrossRef](#)]
61. Lomonaco, V.; Maltoni, D. Comparing Incremental Learning Strategies for Convolutional Neural Networks. In *Artificial Neural Networks in Pattern Recognition*; Lecture Notes in Computer Science; Schwenker, F., Abbas, H.M., El Gayar, N., Trentin, E., Eds.; Springer International Publishing: Cham, Switzerland, 2016; Volume 9896, pp. 175–184. [[CrossRef](#)]