

Article

VRBagged-Net: Ensemble Based Deep Learning Model for Disaster Event Classification

Muhammad Hanif *, Muhammad Atif Tahir  and Muhammad Rafi 

FAST School of Computing, National University of Computer and Emerging Sciences, Karachi 75030, Pakistan; atif.tahir@nu.edu.pk (M.A.T.); muhammad.rafi@nu.edu.pk (M.R.)

* Correspondence: hanif.soomro@nu.edu.pk

Abstract: A flood is an overflow of water that swamps dry land. The gravest effects of flooding are the loss of human life and economic losses. An early warning of these events can be very effective in minimizing the losses. Social media websites such as Twitter and Facebook are quite effective in the efficient dissemination of information pertinent to any emergency. Users on these social networking sites share both textual and rich content images and videos. The Multimedia Evaluation Benchmark (MediaEval) offers challenges in the form of shared tasks to develop and evaluate new algorithms, approaches and technologies for explorations and exploitations of multimedia in decision making for real time problems. Since 2015, the MediaEval has been running a shared task of predicting several aspects of flooding and through these shared tasks, many improvements have been observed. In this paper, the classification framework VRBagged-Net is proposed and implemented for flood classification. The framework utilizes the deep learning models Visual Geometry Group (VGG) and Residual Network (ResNet), along with the technique of Bootstrap aggregating (Bagging). Various disaster-based datasets were selected for the validation of the VRBagged-Net framework. All the datasets belong to the MediaEval Benchmark Workshop, this includes Disaster Image Retrieval from Social Media (DIRSM), Flood Classification for Social Multimedia (FCSM) and Image based News Topic Disambiguation (INTD). VRBagged-Net performed encouraging well in all these datasets with slightly different but relevant tasks. It produces Mean Average Precision at different levels of 98.12, and Average Precision at 480 of 93.64 on DIRSM. On the FCSM dataset, it produces an F1 score of 90.58. Moreover, the framework has been applied on the dataset of Image-Based News Topic Disambiguation (INTD), and exceeds the previous best result by producing an F1 evaluation of 93.76. The VRBagged-Net with a slight modification also ranked first in the flood-related Multimedia Task at the MediaEval Workshop 2020.



Citation: Hanif, M.; Tahir, M.A.; Rafi, M. VRBagged-Net: Ensemble Based Deep Learning Model for Disaster Event Classification. *Electronics* **2021**, *10*, 1411. <https://doi.org/10.3390/electronics10121411>

Academic Editors: Jungong Han and Guiguang Ding

Received: 30 April 2021

Accepted: 8 June 2021

Published: 11 June 2021

Keywords: flood classification; disaster response; road passability; social media

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

Water is one of the most significant substances on earth. A large amount of water is wasted by various means, particularly through the overflow of rivers, lakes and similar streams. As a result, dry land receives an extensive flow of water, which creates flooding. Mostly, flood causes the loss of human lives and severe destruction of valuable resources. An appropriate response to flooding requires a timely and accurate flow of information from the affected area to the responsible organizations. Flood response systems have been improved from manual record keeping, to sensor-based monitoring of sensitive zones. However, sensor-based complex systems also require an enormous quantity of human and hardware resources, to monitor and report flooding conditions [1].

Recently, use of social media has increased greatly, which has generated a lot of valuable data. The enormous amount of social media data is the key to resolve various challenging problems, specifically the monitoring of disastrous situations [2–4]. Disaster-related situations require timely information, so that the damage can be minimized through

appropriate measures from the responsible authorities. As it is very difficult to physically collect information from all flood-sensitive zones, social media data play a significant role in providing valuable and timely information. Different social networks, including twitter and Flickr, provide facilities of uploading and sharing text and images to its subscribers. This enormous quantity of data can be processed to extract useful information to obtain better solutions of various problems including floods.

The focus of using social media data for flooding events has received more attention due to the MediaEval benchmark workshop [5–7]. The workshop aimed to encourage social media and satellite data based solutions for different challenges of the flooding event. MediaEval, 2017 targeted solutions of classifying flooding events with the help of text and images retrieved from social media data [5]. MediaEval, 2018 has focused on the presence of roads and their passability status by using social media data [6]. MediaEval, 2019 evaluated the severity of the flooding situation by predicting a person in an image experiencing a water level above their knees, and the prediction of whether or not an online article is related to a flooding situation [7].

Deep Learning based approaches are widely used for flood classification using social media [4,8–12]. A significant effort has been invested in searching for better solutions using deep learning approaches. Ensemble techniques such as bagging are becoming increasingly significant as they have frequently shown the ability to improve upon the generalization ability of a single deep learning model [3,4,11]. This paper describes our ensemble-based deep learning system VRBaggedNet, presented in MediaEval competitions for flood classification using social media. This system ranked first in the Medieval 2020 flood classification task when evaluated independently by the organizers. The main novelty in this paper is the ensemble of several bagged based Visual Geometry Group (VGG) and Residual Network (ResNet) models. Several deep learning models are trained using Bagging, i.e., by sampling with replacement of training data. These models are later combined using aggregation in order to reduce the over-fitting and the error rate of individual learner. In addition, the proposed ensemble-based system also alleviates the inherent problem of class imbalance in some MediaEval tasks. The following are our main contributions:

- A comprehensive literature survey of the state-of-the-art methods for the classification of different flooding events is provided in this paper. The discussion includes shallow learning- as well as deep learning-based methods for feature extraction and classification of flooding events.
- A VRBagged-Net ensemble classification framework has been proposed for the successful classification of flooding events. Several accurate and diverse models were trained using VGG and ResNet-based deep learning models. These models were then combined using Majority Voting.
- Experiments were conducted on several standard benchmarks for flood classification and compared with the state-of-the-art approaches.

This paper is organized as follows. Section 2 reviews existing work followed by description of data sets in Section 4. Section 3 discusses the proposed methodology with experiments and results in Section 5. Section 6 concludes the paper. The ensemble-based deep learning models have provided various useful outcomes in prediction tasks in different research fields including cancer prediction, speech recognition and crude oil price prediction. Stacking-based ensemble learning has been performed for speech recognition, in which various deep learning neural networks are stacked, including DNN, CNN, and RNN [13]. Researchers have combined the merits of Stacked Denoising Autoencoder (SDAE) and Bootstrap Aggregation (Bagging) and formulated a method for prediction of crude oil price [14]. Another ensemble-based approach has been proposed for the prediction of cancer patients, which has combined different machine learning and deep learning models [15]. Ensemble-based framework IBaggedFCNet has been proposed for the detection of anomalies in videos, which has utilized Inception-v3 and a 3-Layer Fully Connected (FC) neural network [16].

Social media data can also produce fruitful information for disaster response when combined with various other data resources, including hydrological data. An effective multimodal neural network has been designed by combining the text of tweets and hydrological information based on the timestamp and location mentioned in tweets [17]. Another research effort has generated a method for improving situation awareness during disasters by combining social media data with hydrological and sensor-based data [18]. Another research attempt has combined crowdsourced photos and volunteered geographic data to produce an effective method for the estimation of flooding events and identification of its affected regions [19].

2. Literature Review

In recent times, social media data has been used to identify clues of disastrous events. This section of the literature review discusses methods used for the classification of various flooding events. The first part includes the literature, regarding use of social media data for the detection of availability or unavailability of flood in images. The second part of the section includes the literature on identifying the availability of passable roads in the images and the final part discusses methods used to find flood-related topics of articles, from their respective images.

At the MediaEval 2017 workshop, a specific task was designed to combine social media text and images, along with satellite images, to identify flooding events for emergency response [5]. The subtask of the MediaEval, 2017 named as Disaster Image Retrieval from Social Media (DIRSM) [5], provided a dataset of images, along with their relevant text taken from different social media networks. The participants of the workshop used diverse approaches to solve the challenge. Researchers [8] have used Deep Convolutional Neural Network (DCNN) to perform binary image classification on the basis of availability or unavailability of flood. Features have been extracted from images with the help of GoogleNet [9], pretrained on places205, and then extracted features were merged with conventional features, which includes AutoColor Correlation (AC), Edge Histogram (EH), and Tamura. Finally, a Support Vector Machine (SVM) classifier was used to perform binary classification. Another team of researchers [9] used the AlexNet model, pretrained on Places and ImageNet [20] datasets for feature extraction and a Support Vector Machine (SVM) was used for the classification of images [9].

Many other researchers have applied different state-of-the-art methods over the dataset of DIRSM [5]. Researchers [2] have used Spectral Regression in combination with Kernel Discriminant Analysis (SRKDA), over the ensemble of conventional features to predict confidence for binary image classification on the DIRSM dataset. The MultiBrasil team [21] utilized GoogleNet [22], which is pre-trained on the ImageNet dataset, and performed binary classification on the DIRSM dataset. In another research effort [23], X-ResNet [24] pretrained on DeepSentiBank [25] was used, along with Support Vector Machine (SVM) to conduct binary classification of the images of DIRSM dataset. X-ResNet [24] is the extension of ResNet [26].

The prompt response to disastrous events heavily depends upon the availability of passable roads, particularly in a flooding situation. The research was initiated by MediaEval, 2018 [6], which aimed to find the availability of evidence of passable roads. Each instance of dataset comprises of a tweet, with both text and images. Classification was performed by using text, images and the combination of both [6]. Researchers have used text to find the status of roads, but the literature shows that the text-based part of a dataset does not produced any significant outcome, which could help in either finding evidence or status of passable roads [3,10,11,27]. However, another dimension of research has been explored by using image instances given in tweets. Research efforts have utilized various feature extraction and classification techniques to find the status of roads. Different pretrained networks including VGG [28], DenseNet201 [29], Inception V3 [30] and ResNet50 [26] have been used in a variety of methods and obtained promising results [3,10,11,27]. It

has been observed that in comparison to text, visual data have provided significantly better outcomes.

The response system for a flooding event greatly depends on knowledge of the severity of the situation. Medieval Benchmark Workshop 2019 released the dataset for “Multimodal Flood Level Estimation” [7]. The dataset consists of images related to flooding disasters. The dataset was designed to create a better method of image classification on the basis of whether or not one person is available in the image in a standing position, who has water level above their knees [7]. Researchers have utilized different techniques for the detection of such a person and water level above their knees, which includes the use of a 22 layer GoogleNet [22], five fold cross-validation approach with VGG16 and Inception V3, combination of Faster-RCNN, VGG16 and ResNet50 architecture and so on [4,12,31]. In another task of MediaEval Workshop 2019, articles have been analyzed by their images and classified on the basis of whether or not the topic of the article is related to a flooding event [7]. The task is named as “Image-based News Topic Disambiguation” [7]. Prominent research efforts have given more attention to pretrained networks for extraction of features and classification. Few research efforts include utilization of a cross-validation-based approach along with VGG16 and Inception V3 pretrained networks, implementation of the ensemble method by using VGG16 pretrained on ImageNet and Places365 datasets. State-of-the-art results have been evaluated by using the F1 measure and it has been shown that deep learning-based methods have produced highly successful outcomes [4,12,31]. Different ensemble-based approaches have been used in disaster response systems that exploit social media data. The majority of the approaches have utilized the weights of different pre-trained Convolutional Neural Networks (CNN) and various ensemble-based methods. Major challenges faced by the researchers include class imbalance, and use of pre-trained weights from ImageNet dataset, which primarily focuses on object-level information, rather than scene-level information.

3. Methodology

The proposed VRBagged-Net classification framework for flood classification is shown in Figure 1. The framework is divided into three different phases, including preprocessing and stratified random sample selection, transfer learning-based training and prediction of test data. Various components are described as follows:

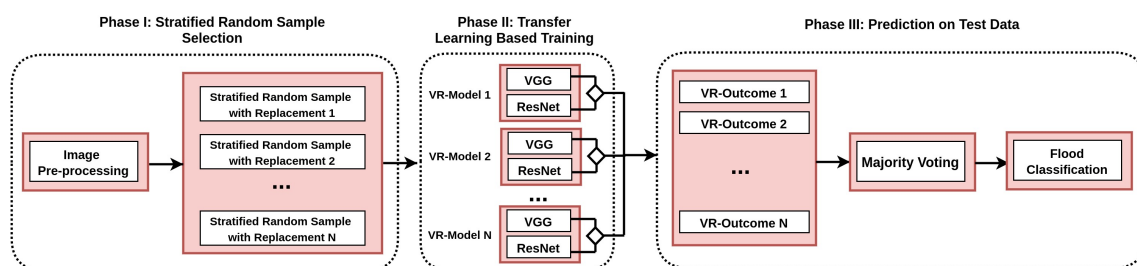


Figure 1. Proposed Framework of VRBagged-Net.

3.1. Preprocessing and Random Sample Selection

In the initial phase, the tasks of preprocessing and stratified random sampling are performed.

3.1.1. Preprocessing

During the preprocessing, images are loaded and resized to a unified dimension. During resizing, all of the images are scaled to have the same heights and widths. Generally, images vary in their height and width, which creates difficulty in learning of classifiers. Hence, images are resized, so that classifiers could be facilitated by similar scale images. During the resizing process, the value of the height and depth of each image was set as 224 pixels.

3.1.2. Stratified Random Sample Selection

After preprocessing, stratified random sampling was applied on selected instances. Stratified random sampling is a probability-based sampling method, in which the population is divided into different sub-groups, called “Strata” (Singular stratum). The formation of an individual stratum is based on the shared characteristics by the members of population. The size of samples from each stratum may be selected by using proportional or equalized methods. In the proportional method, random samples are selected from each stratum according to the size of its respective stratum. However, in equal size sample selection, a fixed quantity of samples is selected from each of the stratum, irrespective of the size of stratum. Selection of equal size instances from each stratum is more efficient, as it equalizes the representation of each stratum [32–35].

Random sampling with replacement is a powerful technique, which utilizes sampling operations to acquire optimal gain [36]. It involves the selection of independent sub-groups from the given population, where the selection of one sub-group does not affect the selection of any other sub-group. The process is used to choose random elements from the population for the creation of one sub-group, and then return those elements to the population. So, the selection of subsequent sub-groups may include previously selected elements. However, there is a chance that few of the elements of the population may not be part of any sub-group. The process is repeated until the required number of sub-groups are generated. In stratified random sampling with replacement, the covariance between different sub-groups is zero, as they are independent of each other. Generally, stratified random sampling with replacement creates effective and reliable representation of each group [37,38].

Therefore, stratified random sampling with replacement is utilized to construct a series of diverse training sets for the flood classification dataset. The population set (S) is super set and used to create multiple subsets S'_1, S'_2, \dots, S'_N .

$$\forall K \mid K \in (1, 15) \ S \supset \{S'_1, S'_2, S'_3, \dots, S'_K\} \quad (1)$$

where (K) is representation of the total quantity of created sub-sets, and we have used a quantity of 15 random samples in VRBagged-Net. However, each of the created sub-sets consist of selected random samples (x') and their relevant labels (y').

$$\forall i \mid i \in K \ S'_i(x'_i, y'_i) \subset S(x, y) \quad (2)$$

where n represents the quantity of instances available in an individual sub-set.

3.2. Transfer Learning Based Training

During the Training phase, transfer learning is implemented to gain advantage of previously learned Convolutional Neural Network architectures. This phase receives each of the training samples generated during phase I, and creates its respective training model. Each of the training models is denoted with the term VR-Model.

VR-Model

VR-Model is the collection of two Convolutional Neural Network architectures, Visual Geometry Group (VGG) [39] and ResNet [26]. A single VR-Model receives its respective stratified random sample with replacement. This individual random sample is utilized by the VR-Model, to train its VGG architecture as well as the ResNet architecture. The versatile combination of VGG and ResNet has created a strong VR-Model.

Visual Geometry Group (VGG16) is a Convolutional Neural Network (CNN) proposed during the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC-2014), which is an annual computer vision competition. The network has secured first and second positions in image localization and image classification tasks, respectively [39]. The VGG-16 (Visual Geometry Group) model is used as a part of the VR-Model. It is pre-trained on both of the

ImageNet [20] and places365 [28] datasets. The ImageNet dataset [20] is helpful in finding object-level information, as it contains more than a million images, which are distributed in 1000 classes. While, the dataset of Places365 helps in recognizing scene-level information and contains about 1.8 million images, distributed in 365 classes. By combination of both ImageNet and Places365 datasets, a hybrid dataset is created, which has 1365 classes. Furthermore, different parameter settings are tried for the architecture of VGG. The last four layers of the network are fine-tuned and remaining layers remain unchanged. The architecture is compiled with the Adam Optimizer and the dropout is set as 0.2. The loss function is used as binary cross entropy and learning rate is set as 10^{-5} . Moreover, the model is trained for 40 epochs. The best performing parameters for the architecture are summarized in Table 1.

Another Convolutional Neural Network used in the VR-Model is ResNet50 (Residual Network) [26]. The selected architecture is a variant of the ResNet model, and pretrained on the ImageNet dataset [20]. Residual Networks (ResNet) has implemented the concept of skip connections, by which various layers of a neural network are skipped during the training of the neural network. Consequently, the problem of vanishing gradient is avoided, and the network can be trained up to a higher number of layers. The model of ResNet accomplished a 3.57% error rate on ImageNet and achieved 1st Place in the classification task of the ILSVRC 2015 [26]. The last eight layers of the architecture are fine-tuned and remaining layers remain unchanged. The architecture is compiled with the Adam Optimizer and the dropout is set as 0.3. Moreover, the loss function is used as a binary cross entropy and the learning rate is set as 10^{-6} . The model is trained for 50 epochs. The best performing parameters for the ResNet50 are summarized in Table 1.

Table 1. Details of CNN-architectures utilized in VR-Model

Parameters	VGG	ResNet
Model	VGG16	ResNet50
Weights	Imagenet + Places365	ImageNet
Epochs	40	50
Learning Rate	10^{-5}	10^{-6}

Each of the VR-Models choose individual sub-set S' , for the training of both of its Convolutional Neural Networks (V and R).

$$(V, R) := S'_1(x'_1, y'_1) \quad (3)$$

3.3. Prediction on Test Data

Final phase of the VRBagged-Net manages predictions through trained VR-Models on test data and ensemble of predictions by different VR-Models. Each of the VR-Models consists of two architectures of Convolutional Neural Networks, Visual Geometry Group (VGG) [39] and ResNet [26], and each of them is separately used for prediction of test instances. In the last layer of each model, the sigmoid activation function is used to retrieve a probabilistic outcome, ranging between 0 and 1. Two Convolutional Neural Networks, VGG and Resnet are used to extract the output, which can be represented by the following:

$$R_{out} = \frac{1}{1 + e^{-R_{pre}}} \quad (4)$$

$$V_{out} = \frac{1}{1 + e^{-V_{pre}}} \quad (5)$$

where R_{pre} and V_{pre} represent the outcome of individual models before implementation of the Sigmoid function. However, R_{out} and V_{out} represent the outcomes retrieved by individual CNN, after the sigmoid function is applied. R_{out} and V_{out} are retrieved in the form of probabilities, ranging from 0 to 1. These probabilities are averaged to combine the prediction by both of the architectures, and represented by \tilde{x}_{VR} .

$$\forall i \mid i \in (1, 15) \tilde{x}_{VR_i} = (V_{out,i} + R_{out,i})/2 \quad (6)$$

Furthermore, the probabilistic averaged outcome \tilde{x}_{VR} is converted into discrete form $L\tilde{x}_{VR}$, as if value of \tilde{x}_{VR} is less than 0.5, then it is considered as 0, otherwise 1.

$$f(L\tilde{x}_{VR_i}) = \begin{cases} 1 & \text{if } \tilde{x}_{VR} \geq 0.5 \\ 0 & \text{if } \tilde{x}_{VR} < 0.5 \end{cases} \quad (7)$$

The steps involved in processing of a single input image, through VRBagged-Net are shown in Figure 2.

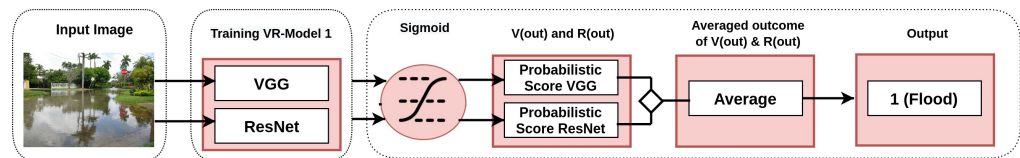


Figure 2. Processing steps of VRBagged-Net for an individual input image.

Subsequently, an ensemble is created by combining discrete predictions of individual VR-Outcome ($L\tilde{x}_{VR}$) and majority voting is applied.

$$\forall i \mid i \in (1, 15) f(y) = \operatorname{argmax}(L\tilde{x}_{VR}) \quad (8)$$

The majority voting yields the final outcome for flood classification, as shown in Figure 1. The ensemble of various VR-Models is a very significant part of VRBagged-Net. The ensemble of versatile VR-Models reduces the error rate of individual classifiers and improves the classification accuracy. The sequence of steps followed in VRBagged-Net is given in Algorithm 1.

Algorithm 1: VRBaggedNet

Input : Dataset[Test,Train] = Train(x,y), Test(xt,yt), K=15
Output: \hat{y} , error

```

1 for  $i = 0 \in K$  do
2    $Sample(x_i, y_i) = Subset(Train_{(x,y)})$  //Training Phase
3    $V_i := Sample_{(x_i, y_i)}$   $R_i := Sample_{(x_i, y_i)}$ 
4    $Weights_{V_i} := VGG_{V_i}$  //Weights are learned from training
   phase
5    $Weights_{R_i} := ResNet_{R_i}$ 
6    $V_{pre,i} := VGG(test_x, W_{V_i})$  //Testing Phase
7    $R_{pre,i} := ResNet(test_x, W_{R_i})$ 
8    $V_{out,i} := \frac{1}{1+e^{-V_{pre,i}}}$ 
9    $R_{out,i} := \frac{1}{1+e^{-V_{pre,i}}}$ 
10   $\tilde{x}_{VRi} = (V_{out,i} + R_{out,i})/2$ 
11  if  $L\tilde{x}_{VRi} = 0.5$  then
12     $L\tilde{x}_{VRi} = 1$ 
13  else
14     $L\tilde{x}_{VRi} = 0$ 
15  end if
16 end for
17  $\hat{y} = argmax(L\tilde{x}_{VRi})$ 
18  $error = \sum_{i=1}^m |y_t - \hat{y}|/m$  //m is number of test samples

```

4. Datasets for Disaster Response System

Various datasets are selected to verify the effectiveness of the VRBagged-Net classification framework. All the selected datasets are related to different flooding events, which include Disaster Image Retrieval from Social Media (DIRSM) Dataset, Flood Classification for Social Multimedia (FCSM) and Image-based News Topic Disambiguation (INTD).

4.1. Dataset 1: Disaster Image Retrieval from Social Media (Dirsm)

The dataset is taken from the task of Disaster Image Retrieval from Social Media dataset (DIRSM), which is a sub-task of The Multimedia Satellite Task at MediaEval Benchmark Workshop 2017 [5]. It includes images taken from different social networks including Flickr and Twitter. All the images, having evidence of flood are kept in class “Flood”, while the images having no any evidence of flood are placed in class “No flood”. In the dataset of 6600 images, 5280 images are part of the train-set and the remaining 1320 are part of the test-set. Figure 3 shows the sample images for the Dataset of DIRSM.

Class: Flood**Class: No Flood**

Figure 3. DIRSM Dataset: Random images of “Flood” and “No flood” classes.

4.2. Dataset 2: Flood Classification for Social Multimedia (FcsM)

The dataset is taken from the task of “Flood Classification for Social Multimedia (FCSM)”, which is sub-task of The Multimedia Satellite Task at MediaEval Benchmark Workshop 2018 [6]. The first objective of the sub-task involves retrieval of images, which provide direct evidence regarding road passability by conventional means including boats, Hummer, Landrover, Monster Trucks and similar vehicles. The FCSM dataset includes 8844 tweets, and each of the tweets contain its relevant image. A quantity of 5818 images are included in the train-set and the remaining 3026 images are part of the test-set. The training set includes 2128 images with direct evidence of road passability and 3685 images with no-direct evidence of road passability. Figure 4 shows the sample images for the FCSM Dataset.

Class: Direct evidence of road pass-ability



Class: No direct evidence of road pass-ability

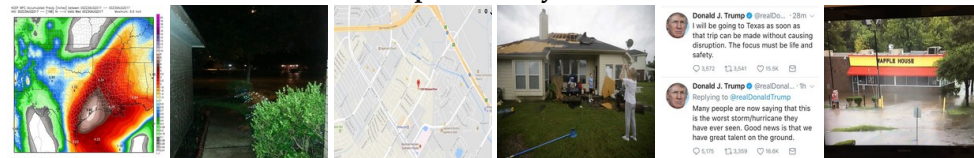


Figure 4. FCSM Dataset: Random images of “Route” and “No Route” classes.

4.3. Dataset 3: Image-Based News Topic Disambiguation (Intd)

MediaEval Benchmark Workshop, 2019 has released a dataset of Image-based News Topic Disambiguation (INTD) [7], which includes a set of images that appeared in online articles. The dataset aimed to predict whether or not the topic of article in which each image appeared was related to the flooding disaster. Topics of the articles of the dataset are ambiguous, even if each of the articles contain related keywords such as “flood”. The development set of the dataset includes 5181 images, while its test set comprises of 1296 images [7]. Figure 5 shows the sample images for the INTD Dataset.

Class: Flood-Related Topic



Class: Not Flood Related Topic



Figure 5. INTD Dataset: Random images of “Flood-Related Topic” and “Not Flood-Related Topic” classes.

4.4. Dataset 4: Flood-Related Multimedia Task at Mediaeval Workshop 2020 (Frmt)

MediaEval Benchmark Workshop, 2020 has released a dataset of Flood-related Multimedia Task at MediaEval workshop 2020 (FRMT) [40], which includes a set of images that appeared in tweets published in Italian. The dataset of the task comprises of 5419 Tweets for (development-set) and 2358 Tweet for test-set. The images were collected from Twitter between 2017 and 2019.

5. Experimental Results and Comprehensive Analysis

The proposed method has been applied on different datasets including DIRSM, FCSM and INTD. The Google Colaboratory (Colab) service has been utilized for the experiments. Colab is a cloud-based service offered by Google and provides support of various artificial intelligence-based libraries as TensorFlow and Keras. The experiments were performed by using the Tesla T4 Graphics Processing Unit (GPU) and the memory of 12 GB. Average training time required for the training of batch containing 32 images, VGG and ResNet require 147 milliseconds and 184 milliseconds, respectively. However, for bagging, all models can be run in parallel. The proposed method has utilized 15 bags, and all these models can be executed in parallel. Moreover, on average, VGG requires 50.73 milliseconds for the prediction of a single image, while ResNet50 needs 43.88 milliseconds to predict a single image. The proposed VRBagged-Net classification framework is applied on the DIRSM. The evaluation of extracted results and their comparison with state-of-the-art results are discussed below:

5.1. Results for Dirsm Dataset

For the DIRSM dataset [5], two types of evaluation measures are used for binary image classification such that Average Precision (APK) at different cut-offs (50, 100, 150, 240 and 480) and Mean Average precision (MAP). The average precision (AP) is relative to the area under a precision–recall curve. It is used to evaluate the average of the precision over images, which are considered relevant. This metric combines precision and recall into a single performance value [41]. AP is defined as:

$$AP = \frac{1}{|N|} \sum_{E=1}^{|N|} C_E \quad (9)$$

where N represents quantity of positive samples in a test set. Moreover, the contribution C_E of the E^{th} element in the ranking list is defined as follows:

$$C_E = \begin{cases} \frac{|N \cap L_E|}{E} & \text{if actual and predicted class true} \\ 0 & \text{if actual and predicted class not true} \end{cases} \quad (10)$$

where $L_E = \{I_1, I_2, \dots, I_E\}$ is a ranked list of the top E retrieved samples from the test set.

The proposed VRBagged-Net classification framework is applied on the DIRSM [5] dataset. Initially, 15 models are selected, each of them containing 1500 images from class 1 (Flood) and class 0 (NoFlood). Each of the selected samples is used for the training of separate VR-Model, subsequently 15 VR-outcomes are generated. Finally, majority voting is applied to retrieve final predictions. The results are evaluated by average precision at k , and achieved the average precision of 100, 100, 100, 98.85 and 93.64, at intervals of 50, 100, 240 and 480, respectively. Moreover, another evaluation measure of mean average precision on different intervals of 50, 100, 150, 240 and 480 revealed 98.12. The top 12 images retrieved by VRBagged-Net for the Dataset of DIRSM are shown in Figure 6.



Figure 6. Top 12 images correctly retrieved by VRBagged-Net for the dataset of DIRSM.

It has been observed that the previous methods have utilized features extracted from various Convolutional Neural Networks (CNN) and combined them; however, the challenge of class-imbalance is not focused. The proposed VRBagged-Net has provided a better outcome, mainly because of balanced random sample with replacement and creation of multiple training models. The minority class of Flood and Noflood of the DIRSM [5] dataset, are equalized by random sample selection with replacement. It is observed by the increasing number of VR-Models, that the performance of the VRBagged-Net framework has also increased. Comparison of VRBagged-Net with state-of-art methods is given in Table 2.

Table 2. DIRSM Dataset: Comparison of VRBagged-Net with state-of-the-art methods.

Reference	Method	AP@50	AP@100	AP@250	AP@480	MAP
Benjamin et al. [23]	X-ResNet	-	-	-	86.64	95.71
Sheharyar et al. [9]	AlexNet	-	-	-	86.81	95.73
Icaro et al. [42]	ResNet, VGG NASNet	100.00	100.00	98.55	88.41	96.74
VRBagged-Net	Bagging, VGG, ResNet	100.00	100.00	98.85	93.64	98.12

5.2. Results for Fcsm Dataset

For the FCSM dataset [6], the F_1 measure has been utilized to evaluate the classification of evidence of road passability and status of passable or not-passable roads. The total VR-Models are trained and their respective outcomes are generated. Finally, majority voting is applied. The implementation of VRBagged-Net revealed F_1 measure of 90.58 for the dataset of FCSM. The top 10 images retrieved by VRBagged-Net for the dataset of FCSM are shown in Figure 7.

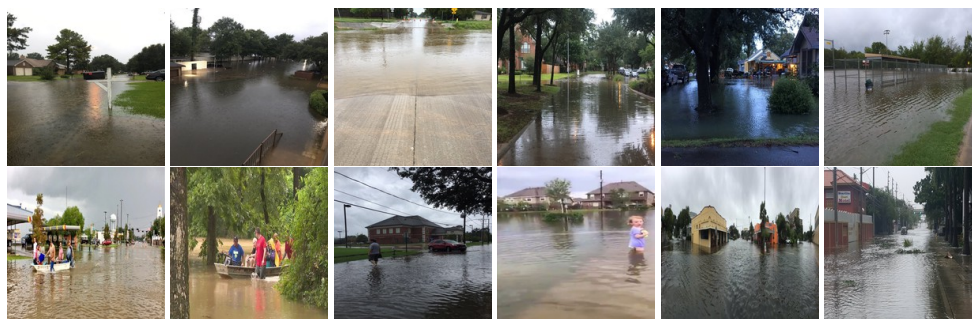


Figure 7. Top 12 images correctly retrieved by VRBagged-Net for the dataset of FCSM.

In the previous best-performing methods, the challenge of finding evidence of road in images has been attempted by combining different pre-trained Convolutional Neural Networks (CNNs). Researchers have mainly utilized various deep learning models including ResNet, VGG, Xception, and DenseNet for feature extraction and classification tasks [3,27,43,44]. Table 3 shows the comparison between the results produced by VRBagged-Net and state-of-the-art methods.

Table 3. FCSM Dataset: Comparison of VRBagged-Net with state-of-art methods

Reference	Method	F1(%)
Hanif et al. [3]	Deep Local Features and SRKDA	74.58
Laura et al. [43]	InceptionV3, VGG, DenseNet, NASNetLarge	87.32
Zhengyu et al. [27]	ResNet50, Support Vector Machine (SVM)	87.58
Benjamin et al. [44]	SVM, Faster R-CNN	87.70
VRBagged-Net	Bagging, VGG, ResNet	90.58

5.3. Results for Intd Dataset

The dataset of INTD [7] involves the challenge of binary classification of images, on the basis of the topic of their respective articles. It is challenging to find topic relevancy of article with its image, as a few of the images are not showing any indication of flood, but the topic of their article is related to flood. However, a few of the images are showing flooding conditions, but the topic of their article is not related to flood. For the INTD dataset [7], VRBagged-Net is applied. As the dataset is highly imbalanced, and ratio of positive class (Flood related topic) to negative class (Not flood related topic) is (1:8). Therefore, different balancing techniques are used, including weight balancing and data augmentation. It is revealed that over-sampling based on data augmentation has produced a better outcome as compared to weight balancing. Image augmentation is a technique of artificially creating new samples of original dataset images, while preserving the labels. Augmentation creates multiple copies of images by using different alterations in images including rotation, elastic deformation and adding noise. It is a very useful technique for expanding the size of the dataset to facilitate the training of deep learning models. For the minority class of INTD [7], the dataset augmentation technique is implemented with the help of python based library, Augmentor [45]. The library utilizes a stochastic, pipeline-based method for the generation of multiple copies of the given images, by using the techniques of rotation, flip and zoom, as shown in Figure 8.

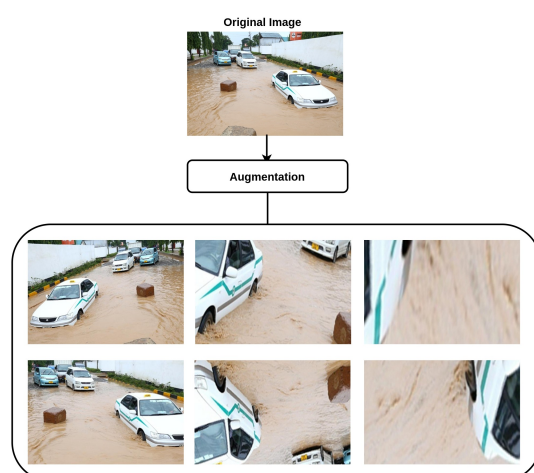


Figure 8. Different augmented samples created from single image.

After the implementation of augmentation, 15 random samples with replacement are created and used for the training of VR-Models. Subsequently, VR-Outcomes are generated by utilizing each of the VR-Models. Finally, majority voting is applied for the final prediction, as shown in Figure 1. A few of the images retrieved by VRBagged-Net for the INTD dataset are shown in Figure 9.



Figure 9. Few of the correctly retrieved images by VRBagged-Net for the INTD dataset.

Researchers [12] have balanced the classes by using the technique of oversampling and then, the VGG16 pretrained network is utilized for the prediction of the topic of the article; the outcome provided an F1 score of 90.2. Another research effort has used instance weight to balance the classes and fine-tuned VGG16 on the INTD dataset. As a result, the an F1 score of 89.6 has been achieved [31]. However, the proposed VRBagged-Net has utilized data augmentation technique for the oversampling of minority class and after ensemble of just five VR-Models, the achieved results provided an F1 score of 93.76. Various other research efforts have utilized a combination of various pretrained Neural Networks, but mainly challenges of class-imbalance and overfitting has decreased the performance of outcome [46–50]. The comparative analysis of results produced by the proposed method, with state-of-the-art methods is given in Table 4.

Table 4. INTD Dataset: Comparison of VRBagged-Net with state-of-the-art methods

Reference	Method	F1(%)
Mirko et al. [50]	Inception, MobileNet, VGG	66.28
Kashif et al. [49]	Fusion of AlexNet, VggNet and ResNet, SVM	82.63
Hariny et al. [48]	VGG19	84.26
Dan et al. [47]	MobileNet	85.26
Khanh-An et al. [46]	EfficientNet	88.5
Mir et al. [4]	Ensemble, VGG16, Data Augmentation	89.5
Pierrick et al. [31]	InceptionV3, MobileNetV2, VGG16	89.6
Stelios et al. [12]	VGG	90.2
VRBagged-Net	Bagging, VGG, ResNet	93.76

5.4. Results for Flood-Related Multimedia Task at Mediaeval 2020 (Frmt)

The FRMT dataset [40] involves the challenge of binary classification of images published in Italian Tweets, on the basis of availability or unavailability of flooding conditions. For FRMT [40], only VGG16 is used as the base deep learning model. The rest of the network remains the same, as shown in Figure 1. The results produced by the variation of VRBagged-Net ranked top, by revealing the F1 Score as 20.76; the result is already published in MediaEval workshop, 2020 [40]. The comparative analysis of the results produced by the proposed method, with state-of-the-art methods is given in Table 5.

Table 5. FRMT Dataset: Comparison of VRBagged-Net with state-of-the-art methods.

Reference	Method	F1(%)
Naina et al. [51]	DenseNet, VggNet-19, ResNet	12.91
Rabiul et al. [52]	Xception	14.36
Hanif et al. [53] (VRBagged-Net variation)	Ensemble, VGG16	20.76

Researchers [52] have utilized Xception, pretrained on ImageNet dataset and InceptionV3, but features extracted from the Xception pretrained network have produced better results.

Researchers [51] have utilized three different pretrained networks for the extraction of features of individual instance of dataset including DenseNet, VggNet-19, and ResNet. All the selected networks have been pretrained using ImageNet dataset, to extract object level information. Later, the Support Vector Machine (SVM) is applied for the purpose of classification on each of the extracted feature sets. This FRMT dataset contains various challenges, which have negatively affected the results. The major issue includes the text of tweets that has been annotated without considering their relevant images. Manual inspection of the test set revealed that those images which are considered flood relevant do not contain any visual evidence of flooding events. The complexity of the test set has also

been increased due to symbols of meteorological alerts, which are difficult to understand for any algorithm.

6. Conclusions

This research has proposed a Bagging-based deep learning approach with random input samples and has created a VRBagged-Net classification framework. VRBagged-Net is featured by stratified random sample selection with replacement for the generation of various random samples, and training of their respective models. It has been observed that the combination of weights of object-level and scene-level information for a deep learning model are useful for flood classification. This approach has exceeded previous best results for the datasets of Disaster Image Retrieval from Social Media (DIRSM) [5], Flood Classification for Social Multimedia (FCSM) [6], Image-Based News Topic Disambiguation (INTD) [7] and Flood-Related Multimedia Task at MediaEval 2020 (FRMT) [40]. More variations of different ensembling techniques may be implemented in the future, including AdaBoot and Stacking. Pre-trained models of Neural Network on some domain-specific data are available, such as NasNet [54]. Their utilization is still an open research problem. These efforts may improve the performance of the proposed framework.

Abbreviations

The following abbreviations are used in this manuscript:

AC	AutoColor Correlation
APK	Average Precision
CNNs	Convolutional Neural Networks
DCNN	Deep Convolutional Neural Network
DIRSM	Disaster Image Retrieval from Social Media
EH	Edge Histogram
FCSM	Flood Classification for Social Multimedia
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
INTD	Image-Based News Topic Disambiguation
NLP	Natural Language Processing
ResNet	Residual Network
SVM	Support Vector Machine
SRKDA	Spectral Regression in combination with Kernel Discriminant Analysis
TFIDF	Term-Frequency Inverse Document Frequency
VGG	Visual Geometry Group
VRBagged-Net	Visual Geometry Group Residual Bagging Network
FRMT	Flood-related Multimedia Task

Author Contributions: Conceptualization, M.H., M.A.T. and M.R.; methodology, M.H., M.A.T. and M.R.; validation, M.H., M.A.T.; data curation, M.H.; writing—original draft preparation, M.H., M.A.T. and M.R.; writing—review and editing, M.H., M.A.T. and M.R.; visualization, M.H., M.A.T. and M.R.; supervision, M.A.T. and M.R.; funding acquisition, M.A.T. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Higher Education Commission (HEC) Pakistan and Ministry of Planning Development and Reforms under the National Center in Big Data and Cloud Computing.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hughes, D.; Greenwood, P.; Blair, G.; Coulson, G.; Pappenberger, F.; Smith, P.; Beven, K. An intelligent and adaptable grid-based flood monitoring and warning system. In Proceedings of the 2006 UK eScience All Hands Meeting, Nottingham, UK, 18–21 September 2006.
2. Hanif, M.; Tahir, M.A.; Khan, M.; Rafi, M. Flood detection using social media data and spectral regression based kernel discriminant analysis. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.

3. Hanif, M.; Tahir, M.A.; Rafi, M. Detection of passable roads using ensemble of global and local. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
4. Mir Murtaza, M.H.; Tahir, M.A.; Rafi, M. Ensemble and inference based methods for flood severity estimation using visual data. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
5. Bischke, B.; Helber, P.; Schulze, C.; Venkat, S.; Dengel, A.; Borth, D. The multimedia satellite task at mediaeval 2017: Emergence response for flooding events. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.
6. Benjamin, B.; Patrick, H.; Zhengyu, Z.; Damian, B. The multimedia satellite task at mediaeval 2018: Emergency response for flooding events. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
7. Bischke, B.; Helber, P.; Brugman, S.; Basar, E.; Zhao, Z.; Larson, M.; Pogorelov, K. The multimedia satellite task at mediaeval 2019: Flood severity estimation. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
8. Avgerinakis, K.; Moutzidou, A.; Andreadis, S.; Michail, E.; Gialampoukidis, I.; Vrochidis, S.; Kompatsiaris, I. Visual and textual analysis of social media and satellite images for flood detection@ multimedia satellite task mediaeval 2017. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.
9. Ahmad, S.; Ahmad, K.; Ahmad, N.; Conci, N. Convolutional neural networks for disaster images retrieval. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.
10. Moutzidou, A.; Giannakeris, P.; Andreadis, S.; Mavropoulos, A.; Meditskos, G.; Gialampoukidis, I.; Avgerinakis, K.; Vrochidis, S.; Kompatsiaris, I. A multimodal approach in estimating road passability through a flooded area using social media and satellite images. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
11. Feng, Y.; Shebotnov, S.; Brenner, C.; Sester, M. Ensembled convolutional neural network models for retrieving flood relevant tweets. *Image* **2018**, *10*, 1.
12. Andreadis, S.; Bakratsas, M.; Giannakeris, P.; Moutzidou, A.; Gialampoukidis, I.; Vrochidis, S.; Kompatsiaris, I. Multimedia analysis techniques for flood detection using images, articles and satellite imagery. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
13. Deng, L.; Platt, J.C. Ensemble deep learning for speech recognition. In Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014.
14. Zhao, Y.; Li, J.; Yu, L. A deep learning ensemble approach for crude oil price forecasting. *Energy Econ.* **2017**, *66*, 9–16. [\[CrossRef\]](#)
15. Xiao, Y.; Wu, J.; Lin, Z.; Zhao, X. A deep learning-based multi-model ensemble method for cancer prediction. *Comput. Methods Progr. Biomed.* **2018**, *153*, 1–9. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Zahid, Y.; Tahir, M.A.; Durrani, N.M.; Bouridane, A. Ibaggedfcnet: An ensemble framework for anomaly detection in surveillance videos. *IEEE Access* **2020**, *8*, 220620–220630. [\[CrossRef\]](#)
17. De Bruijn, J.A.; De Moel, H.; Weerts, A.H.; de Ruiter, M.C.; Basar, E.; Eilander, D.; Aerts, J.C. Improving the classification of flood tweets with contextual hydrological information in a multimodal neural network. *Comput. Geosci.* **2020**, *140*, 104485. [\[CrossRef\]](#)
18. De Albuquerque, J.P.; Herfort, B.; Brenning, A.; Zipf, A. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 667–689. [\[CrossRef\]](#)
19. Schnebele, E.; Cervone, G.; Waters, N. Road assessment after flood events using non-authoritative data. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1007–1015. [\[CrossRef\]](#)
20. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.-F. 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.
21. Nogueira, K.; Fadel, S.G.; Dourado, Í.C.; de Oliveira Werneck, R.; Muñoz, J.A.; Penatti, O.A.; Calumby, R.T.; Li, L.; dos Santos, J.A.; da Silva Torres, R. Data-driven flood detection using neural networks. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.
22. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
23. Bischke, B.; Bhardwaj, P.; Gautam, A.; Helber, P.; Borth, D.; Dengel, A. Detection of flooding events in social multimedia and satellite imagery using deep neural networks. In Proceedings of the MediaEval 2017 Workshop, Dublin, Ireland, 13–15 September 2017.
24. Jou, B.; Chang, S.-F. Deep cross residual learning for multitask visual recognition. In Proceedings of the 24th ACM international conference on Multimedia, Amsterdam, The Netherlands, 15–19 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 998–1007.
25. Chen, T.; Borth, D.; Darrell, T.; Chang, S.-F. DeepSentibank: Visual sentiment concept classification with deep convolutional neural networks. *arXiv* **2014**, arXiv:1410.8586.
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
27. Zhao, Z.; Larson, M.A.; Oostdijk, N.H.J. Exploiting local semantic concepts for flooding-related social image classification. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
28. Zhou, B.; Lapedriza, A.; Khosla, A.; Oliva, A.; Torralba, A. Places: A 10 million image database for scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 1452–1464. [\[CrossRef\]](#) [\[PubMed\]](#)

29. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
30. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
31. Bruneau, P.; Tamisier, T. Transfer learning and mixed input deep neural networks for estimating flood severity in news content. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
32. Neyman, J. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. In *Breakthroughs in Statistics*; Springer: New York, NY, USA, 1992; pp. 123–150.
33. Parsons, V.L. *Stratified Sampling*; American Cancer Society: Atlanta, GA, USA, 2017; pp. 1–11.
34. Lumley, T. *Complex Surveys: A Guide to Analysis Using R*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 565.
35. Aoyama, H. A study of the stratified random sampling. *Ann. Inst. Stat. Math.* **1954**, *6*, 1–36. [\[CrossRef\]](#)
36. Sahu, A.; Runger, G.; Apley, D. Image denoising with a multi-phase kernel principal component approach and an ensemble version. In Proceedings of the 2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 11–13 October 2011; pp. 1–7.
37. Basu, D. On sampling with and without replacement. *Sankhyā Indian J. Stat.* **1958**, *20*, 287–294.
38. Sengupta, S. On comparisons of with and without replacement sampling strategies for estimating finite population mean in randomized response surveys. *Sankhya B* **2016**, *78*, 66–77. [\[CrossRef\]](#)
39. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
40. Andreadis, S.; Gialampoukidis, I.; Karakostas, A.; Vrochidis, S.; Kompatsiaris, I.; Fiorin, R.; Norbiato, D.; Ferri, M. The flood-related multimedia task at mediaeval 2020. In Proceedings of the MediaEval 2020 Workshop, Online, 14–15 December 2020.
41. Tahir, M.A.; Yan, F.; Koniusz, P.; Awais, M.; Barnard, M.; Mikolajczyk, K.; Bouridane, A.; Kittler, J. A robust and scalable visual category and action recognition system using kernel discriminant analysis with spectral regression. *IEEE Trans. Multimed.* **2013**, *15*, 1653–1664. [\[CrossRef\]](#)
42. Dourado, I.C.; Tabbone, S.; da Silva Torres, R. Event Prediction based on Unsupervised Graph-Based Rank-Fusion Models. In *International Workshop on Graph-Based Representations in Pattern Recognition*; Springer International Publishing: Berlin, Germany, 2019; pp. 88–98.
43. Lopez-Fuentes, L.; Farasin, A.; Skinnemoen, H.; Garza, P. Deep Learning Models for Passability Detection of Flooded Roads. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
44. Bischke, B.; Helber, P.; Dengel, A. Global-Local Feature Fusion for Image Classification of Flood Affected Roads from Social Multimedia. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
45. Bloice, M.D.; Stocker, C.; Holzinger, A. Augmentor: An image augmentation library for machine learning. *J. Open Source Softw.* **2017**, *2*, 8. [\[CrossRef\]](#)
46. Quan, K.A.C.; Nguyen, T.C.; Nguyen, V.T.; Tran, M.T. Flood Event Analysis base on Pose Estimation and Water-related Scene Recognition. In Proceedings of the MediaEval 2019 Workshop, SophiaAntipolis, France, 27–29 October 2019.
47. Bînă, D.; Vlad, G.A.; Onose, C.; Cercel, D.C. Flood severity estimation in news articles using deep learning approaches. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
48. Ganapathy, H.; Bandlamudi, G.; Yamini, L.; Bhuvana, J.; Mirnalinee, T.T. Deep learning models for estimation of flood severity using satellite and news article images. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
49. Ahmad, K.; Pogorelov, K.; Ullah, M.; Riegler, M.; Conci, N.; Langguth, J.; Al-Fuqaha, A. Multi-modal machine learning for flood detection in news, social media and satellite sequences. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
50. Zaffaroni, M.; Lopez-Fuentes, L.; Farasin, A.; Garza, P.; Skinnemoen, H. Ai-based flood event understanding and quantification using online media and satellite data. In Proceedings of the MediaEval 2019 Workshop, Sophia Antipolis, France, 27–29 October 2019.
51. Said, N.; Ahmad, K.; Gul, A.; Ahmad, N.; Al-Fuqaha, A. Floods detection in twitter text and images. In Proceedings of the MediaEval 2020 Workshop, Online, 14–15 December 2020.
52. Islam, R.; Alan, W. Flood detection in twitter using a novel learning method for neural networks. In Proceedings of the MediaEval 2020 Workshop, Online, 14–15 December 2020.
53. Hanif, M.; Joozer, M.; Huzaifa, M.; Rafi, M. An ensemble based method for the classification of flooding event using social media data. In Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29–31 October 2018.
54. Zoph, B.; Vasudevan, V.; Shlens, J.; Le, Q.V. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8697–8710.