

# Camera-First Form Filling: Reducing the Friction in Climate Hazard Reporting

Kristina Wolf

k.wolf2@newcastle.ac.uk  
Newcastle University  
Newcastle upon Tyne, UK

Dominik Winecki

winecki.1@osu.edu  
The Ohio State University  
Columbus, Ohio, USA

Arnab Nandi

nandi.9@osu.edu  
The Ohio State University  
Columbus, Ohio, USA

## ABSTRACT

The effective reporting of climate hazards, such as flash floods, hurricanes, and earthquakes, is critical. To quickly and correctly assess the situation and deploy resources, emergency services often rely on citizen reports that must be timely, comprehensive, and accurate. The pervasive availability and use of smartphone cameras allow the transmission of dynamic incident information from citizens in near-real-time. While high-quality reporting is beneficial, generating such reports can place an additional burden on citizens who are already suffering from the stress of a climate-related disaster. Furthermore, reporting methods are often challenging to use, due to their length and complexity. In this paper, we explore reducing the friction of climate hazard reporting by automating parts of the form-filling process. By building on existing computer vision and natural language models, we demonstrate the automated generation of a full-form hazard impact assessment report from a single photograph. Our proposed data pipeline can be integrated with existing systems and used with geospatial data solutions, such as flood hazard maps.

### ACM Reference Format:

Kristina Wolf, Dominik Winecki, and Arnab Nandi. 2023. Camera-First Form Filling: Reducing the Friction in Climate Hazard Reporting. In *Workshop on Human-In-the-Loop Data Analytics (HILDA '23), June 18, 2023, Seattle, WA, USA*. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3597465.3605218>

## 1 INTRODUCTION

The prevalence of natural hazards is likely to increase due to the rise in weather and climate extremes, leading to adverse impacts and risks to human and natural systems worldwide [22]. The United Nations consolidated various measures to combat climate change in their 2030 Agenda for Sustainable Development Goals [15]. In particular, Goal 13 calls for urgent action to improve resilience to natural hazards. Due to increasing climate disasters and the resulting incidents, the demand for, and the pressure on, emergency services is rising [25].

Emergency services use spatial data tools and location-based data from various agencies to understand risks and take preventive

measures, such as detecting weather changes and monitoring sea level rise[5, 30, 31]. Although these tools provide valuable information, they rely on scheduled batch updates and hence do not reflect real-time conditions. Data collected by citizens during natural hazards can offer more immediate and dynamic information, enhancing the situational awareness of emergency services [6].

The pervasiveness of communications infrastructure has enabled many different modes for citizens to contact emergency services in the event of a climate-related hazard. For reporting non-emergency incidents, local city councils and municipalities offer reporting tools that citizens can access via their phones or websites, such as “311” services in Canada [21] and the United States [18]. However, calls to a “311” service do not scale well in emergencies as they require human personnel who are likely to answer emergency calls instead. While phone calls are still the most common communication channel, others are becoming more accessible, such as text messaging, online reporting platforms, and social media. Website-based reporting does not present scale issues, but it can impose other difficulties: citizens must fill out and submit forms which tend to be long and complex. They are asked to type out a description of the incident, indicate the incident’s location on a map, and select the appropriate incident category from a large number of categorical drop-down fields. Since these are usually performed on a mobile device, users are significantly slower [12] when filling in forms due to touchscreens’ limited text input capabilities [23, 26]. In addition to this interaction difficulty, there is often uncertainty about the contents of a report. A citizen may need help understanding jargon or acronyms, the expected format of the reports, or what information is in scope. Reporting detailed information during extreme and rushed circumstances is already a high-friction event for the citizen. The length and complexity of any given reporting method create an inherent trade-off between the number of reports received and the comprehensiveness and accuracy of each report. Longer and highly structured forms reduce data quality since the complexity may lead to submission errors or citizens giving up on the report entirely. Moreover, as non-emergency reports are voluntary, citizens are less likely to accept the cost to receive any potential benefit. As a result, emergency responders receiving the report may not have sufficient high-quality data, leading to the incident not being resolved in the intended manner. Therefore, to better assess the impacts and resulting risks to human and natural systems, we need effective tools and services to lower all possible friction, and enable a timely response from multi-agencies.

**Adoption of AI-based services for emergency services:** Currently available commercial solutions for emergency services include OneCall, a Hexagon solution that can help operators in the command and control rooms to detect complex emergencies sooner

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*HILDA '23, June 18, 2023, Seattle, WA, USA*

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 979-8-4007-0216-7/23/06...\$15.00  
<https://doi.org/10.1145/3597465.3605218>

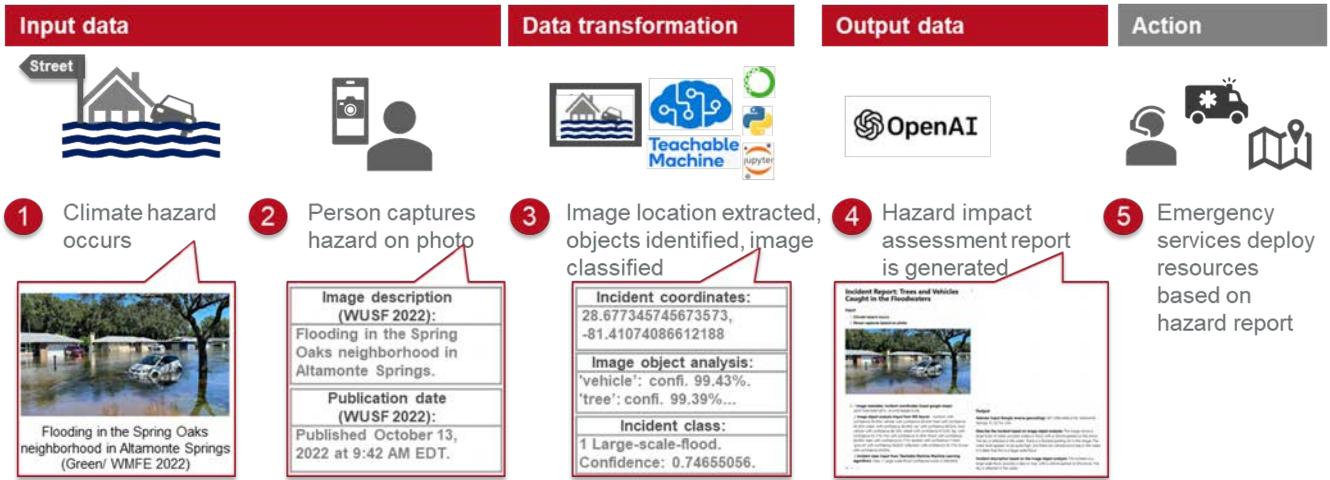


Figure 1: Workflow for a camera-first form filling approach

through assistive integrated AI [9]; RedBox Voice's AI-supported speech analytics to support emergency services in reporting incidents [24]; and Dataminr's First Alert product for the public sector to alert first responders about breaking events and enable a faster response [3]. Despite some early adopters who use the above software, a study conducted by Five9 [13] in the UK highlights the untapped potential of AI services within existing emergency services, such as assisting in handling a large number of calls and collecting incident-relevant information to share amongst responders [1]. While implementing commonly agreed standards helps manage the overall situation in an incident between different groups of responders, there are still common challenges involved in current multi-agency incident response processes, such as [33]:

- Redundancies in communication streams;
- Missing joint overview across agencies;
- Data silos which limit information sharing;
- Heterogeneous IT system landscape; and
- Lack of integrated real-time data on current weather, traffic, and hazard conditions.

Additional hurdles to adoption of AI-based services include availability of funding, manual interaction during multi-agency incident coordination, and a need for more awareness of the potential of using AI [1, 13, 20]. Since adopting new systems can increase the complexity of already fragmented infrastructure, it is important to consider augmenting existing workflows than replacing them.

**Reducing the Friction:** To address these issues, we take the approach of using a *camera-first* approach. Can we reduce the friction of form-filling for the citizen by auto-populating a report based on a photograph taken on their smartphone? This paper explores this question, and presents an approach that builds on publicly available data, models, and tools. As a proof-of-concept, we demonstrate how combining off-the-shelf tools and services including Google Teachable Machine, Microsoft Azure Cognitive Image API, and OpenAI's GPT models can accelerate the process of analysing incoming incident data, extract valuable information from submitted images, and help generate an automated hazard impact report.

With this approach, we aim to reduce friction for citizens in time-critical situations when reporting incidents and support emergency services in their response through (near-) real-time image analysis showing the impact of the hazard. Overall, this study contributes to facilitating multi-agency reporting of climate hazard impacts.

## 2 CASE STUDY

### 2.1 Reducing friction using camera-first reporting

Given the challenges of current approaches, we look towards methods to reduce the friction in climate hazard reporting. This is important for various reasons: First, requiring impacted citizens to fill out complex and arduous forms, especially during emergencies, is an unreasonable ask. Second, citizens may have literacy or language barriers, a hurdle to accurate and complete reporting. Finally, emergency services may require information in a specific format for cross-agency distribution or need particular questions answered, which the impacted citizens may not think to provide. Simultaneously, smartphone adoption is now at an unprecedented level worldwide. Over 6 billion people use smartphones [4] with high penetration numbers across countries irrespective of economic status or urban compared to rural population [16]. Given such high adoption rates, we propose using smartphone cameras as a primary method for hazard reporting, where the impacted citizen takes a photograph, and the remainder of the process is offloaded to our system. Similarly, there has been drastic progress in computer vision and AI models, with profound implications for science and real-world applications. In this study, we combine three complementary concepts to generate structured and comprehensive reports from photographs taken by impacted citizens:

- Supervised image classification;
- Unsupervised image understanding;
- Object recognition; and
- Generative AI models for text generation.

Due to the widespread availability of machine learning frameworks (e.g., PyTorch, TensorFlow), computer vision services from cloud providers (e.g., Google Cloud, AWS, Azure), and pre-trained models (e.g., OpenAI GPT, MobileNet), we observe that it is now possible to build complex AI-based workflows with minimal technical effort or expertise.

## 2.2 Design and development of an automated climate hazard report through computer vision and AI-based text generation

A demonstration workflow for generating a pre-filled report using only a smartphone camera photograph is described herein and visualized in Figure 1. We assume that photographs are geotagged – either the smartphone embeds this information in the EXIF (the standard for metadata in media containers) fields or attaches the coordinates when uploading the photograph. Our proof-of-concept workflow is as follows:

- An image classifier using MobileNet via Google’s no-code *Teachable Machine* tool and a collection of images of different climate hazards. This allows us to classify the incoming images into different categories of natural hazards, such as large-scale or small-scale hazards, and different types, including floods and earthquakes.
- An off-the-shelf Microsoft Azure Cognitive Image API to detect and extract objects from images indicating if urban assets, environmental infrastructure, and people might be located in spatial proximity and could be impacted by the hazard.
- An off-the-shelf Microsoft Azure Cognitive Image API to provide a text description of the image.
- OpenAI’s GPT models (accessed through *Lex.page*) to generate structured text descriptions based on keywords and predicted class labels provided by the Azure Cognitive Image API. This involves providing the model’s content and posing a series of specific questions.

Each of these models, tools, and APIs was selected for their rapid prototyping and generalizability. Google Teachable Machine allowed us to quickly train a custom model which we could deploy ourselves. Microsoft Azure Cognitive Image API is a major cloud model API similar to those offered by other large cloud providers. While Microsoft has some native geospatial capabilities, it is also compatible with different Geographic Information Systems (GIS), such as QGIS and ArcGIS [14]. While there are various ways to store raw data from the computer vision algorithm directly in Microsoft (e.g., Azure SQL DB or Azure Data Storage), we suggest using GIS software to integrate and transform the data and combine it with other geospatial information about the incident area [32]. This integrated picture can improve the situational awareness of several emergency services. OpenAI GPT-3, and later GPT-4, were chosen for their comprehension abilities and ease of prompting. We address architectural considerations in deployments of such models in Section 3.3.

Figure 1 shows the flow of the individual steps from a climate hazard (Hurricane Ian used here) occurs by a citizen who then captures the hazard in a photo which includes meta-information such as an image description, date, time, and the location of the incident.

The next step, data transformation, involves reverse geocoding the incident coordinates, classifying the event type, and recognizing objects. This transformed data is then fed through generative natural language models to create the final report.

## 2.3 Evaluation Considerations

**2.3.1 Evaluation Form.** Hazard reporting forms vary greatly between localities which makes finding a representative form difficult. In support of multi-agency interoperability, the US introduced the National Incident Management System (NIMS), a nationwide, systematic approach to incident management, including the command and coordination of incidents, resource management, and information management [19]. Similarly, the UK adheres to the Joint Emergency Services Interoperability Principles (JESIP) to improve and standardise how the police, fire and rescue, and ambulance services work together when responding to major multi-agency incidents [10]. In this study, we use questions from JESIP M/ETHANE, a standardised reporting framework used by emergency services in the UK to share incident information between responders and control centres [11].

**2.3.2 Evaluation Criteria.** In our use case, the primary criteria of pre-filled form value is the similarity to human input. If the forms resemble human input, the citizen only needs to correct inaccuracies and omissions, rather than inputting the entire form body from scratch. Additionally, the pre-filled form values prime the citizen to respond in a similar format, and, since they are provided with example answers, may be able to better infer the intended questions.

## 2.4 Results: Hazard impact assessment report for Hurricane Ian (USA, September 2022)

We now provide an example of a structured hazard impact assessment report generated from a single image (taken by WUSF National Public Radio) following Hurricane Ian on October 13, 2022 [8]. Hurricane Ian made landfall in Florida in late September 2022 and was classified as a Category 4 storm with sustained winds of 150 mph, making it the fifth strongest storm ever to impact the US [17]. Table 1 lists the inputs corresponding to the data collected from the time the incident was reported through the data transformation and further analysis of the data (see steps 1 to 3 of the workflow as shown in Figure 1). The report sections in *italics*, including the complete sentences, were generated using AI services.

Table 2 provides the impact assessment report (see step 4 of the workflow in Figure 1), which follows the M/ETHANE report structure.

## 3 DISCUSSION

### 3.1 Evaluation of the hazard impact identified by AI on identified objects

While existing M/ETHANE reports provide vital information to the responding agencies, we show how we can further enhance their impact assessment through the analysis of imagery data. AI-assisted image analysis can provide a quicker assessment of the scope of response required, which would have taken more time and effort if responders had extracted and analysed image data manually.

Input	
Image (WUSF 2022):	
Image description (WUSF 2022):	Flooding in the Serenity Oaks neighborhood in Evergreen Springs.
Publication date (WUSF 2022):	Published October 13, 2022 at 9:42 AM EDT
Incident class (Teachable Machine):	<i>Class: 1 Large-scale-flood Confidence score: 0.74655056</i>
Incident coordinates (Google Maps)	28.679345745673573, -81.41474086612188
Address (Google reverse geocoding):	598 Sparrow Ct, Evergreen Springs, FL 32724, USA
Image object analysis (MS Azure):	<p><i>'outdoor' with confidence 99.95%</i>  <i>'vehicle' with confidence 99.43%</i>  <i>'tree' with confidence 99.39%</i>  <i>'water' with confidence 99.09%</i>  <i>'car' with confidence 98.52%</i>  <i>'land vehicle' with confidence 98.18%</i>  <i>'wheel' with confidence 97.20%</i>  <i>'sky' with confidence 92.71%</i>  <i>'tire' with confidence 91.90%</i>  <i>'flood' with confidence 84.66%</i>  <i>'lake' with confidence 81.77%</i>  <i>'parked' with confidence 71.69%</i>  <i>'ground' with confidence 64.82%</i>  <i>'reflection' with confidence 59.17%</i>  <i>'house' with confidence 54.85%</i></p> <p><i>A flooded parking lot.</i></p>
Image description (MS Azure):	<i>A flooded parking lot.</i>

**Table 1: Input data for AI-based text generation (own table, Microsoft Azure Computer Vision, Google Teachable Machine output in blue italics).**

To maximise the time available for emergency services to respond to the incident efficiently and deploy the most suitable response teams, the AI-assisted service helps extract objects visible in the image, such as vehicles, cars, and trees. For detecting and classifying objects in the image, we use Convolutional Neural Networks (CNN), that are pre-trained image classification models from Microsoft. While we use an off-the-shelf approach for our case study, we expect this sub-task to be replicable using the state-of-the-art computer vision and machine learning techniques [29]. As an example, Sarp et al. use a Mask-R-CNN algorithm, a deep learning Region-Based Convolutional Neural Networks (R-CNN) algorithm for object detection and semantic segmentation for detecting floodwater on roadways [27]. Saraza et al. found that a pre-trained deep neural network (VGG-Net) for deep learning-based feature extraction with logistic regression classifier resulted in high scores on all performance measures [28]. For flood area segmentation tasks, they found that the Fully Convolutional Neural Network (FCN) offered the highest score. Similar to these approaches, we expect related models could be applied to identify and label objects, such as floodwater, cars, roads, and buildings, and indicate the color of the identified vehicles.

After determining that floodwater was visible in the image, we tried to refine our questions and extract further information from the identified body of water, such as the depth of the floodwater. The AI-generated output by Lex.page editor was “*quite high*”, as in, “The water level appears to be *quite high*”. Although this automatically generated assessment provides a good first understanding of the flood risk, emergency services require a more accurate assessment,

M	Major Incident	Has a Major Incident been declared? (YES/NO) (If no, then complete M/ETHANE message)	<i>YES - Taking into account the reported incident coordinates and the image object analysis, the affected areas include trees, vehicles and floodwaters.</i>
E	Exact Location	What is the exact location or geographical area of incident?	<i>The exact location of the incident is 28.679345745673573, -81.41474086612188</i>
T	Type of Incident	What kind of incident is it?	<i>The incident is a large-scale flood, as classified by the Teachable Machine Machine Learning algorithm, with a confidence score of 0.74655056.</i>
		What hazards or potential hazards can be identified?	<p><i>The identified hazards or potential hazards include:</i></p> <ul style="list-style-type: none"> <li><i>Floodwater covering a significant portion of the image.</i></li> <li><i>A white car partially submerged in the floodwater.</i></li> <li><i>A red vehicle also partially submerged in the floodwater.</i></li> <li><i>Trees and branches on the left side of the image.</i></li> <li><i>A road or pathway visible on the right side of the image.</i></li> <li><i>A white building or structure in the background on the right side of the image.</i></li> </ul>
		What is the flood depth when putting the water in relation to different identifiable objects:	<i>If we assume that the white car in the center of the image has standard tires with a diameter of around 65 centimeters, and that the water level is at the bottom of the car doors, then the water depth would be approximately 65-70 centimeters.</i>
		What is the flood depth when compared to the door of the car?	<i>Estimate the depth of the flood water through the average size of a car door height, which is typically around 1.4 meters (140 cm). Based on the image, it appears that the floodwater is around halfway up the car door, so we can estimate the depth to be approximately 70 cm.</i>
H	Hazards	What is the flood depth when put in relation to the tires of the car?	<i>Based on the estimated diameter of a car tire, which is around 25 inches or 0.63 meters, and the portion of the tire submerged in the water, it appears that the flood water is at least several inches deep, possibly around 6-8 inches or 15-20 cm.</i>
A	Access	What are the best routes for access and egress?	<i>The best routes of access and egress are: the nearby roads, paths, and waterways that are not blocked by the floodwaters.</i>
N	Number of Casualties	<p>How many casualties are there and what condition are they in?</p> <p>Can you identify people?</p>	<p><i>At this time, there are no reported casualties.</i></p> <p><i>There do not appear to be any visible signs of people in the frame.</i></p>
E	Emergency Services	Which and how many emergency responder assets/personnel are required or are already on-scene?	<i>Emergency responders (firefighters, police officers, or members of a local or national disaster response team) may have been called to the scene to assist with evacuation, rescue operations, or other related activities.</i>

**Table 2: M/ETHANE incident impact report using AI-based text generation in Lex.page AI and OpenAI GPT-3 (own table, AI-text in blue italics).**

e.g., quantifiable water estimates to respond more efficiently. When using OpenAI’s GPT-3, the model explained that it is easier to accurately determine the floodwater depth in the image with additional information or a reference point of known size. However, some objects in the image can help roughly estimate the water depth. When we asked if we could estimate the water depth by identified objects, we received an estimate of 65-70 cm. We then narrowed the question down to two features of the car: *tires* and *doors*. When we related the high water to the door, the estimate of the high water

Has a Major Incident been declared? (YES/NO)	<i>Based on the information provided, it is not clear whether a Major Incident has been declared.</i>
What is the exact location or geographical area of incident?	<i>The exact location of the incident is at 598 Sparrow Ct, Evergreen Springs, FL 32724, USA. The geographical coordinates are 28.679345745673573, -81.41474086612188.</i>
What kind of incident is it?	<i>The incident is a large-scale flood, as indicated by the incident class and the image description from WUSF 2022 which mentions flooding in the Serenity Oaks neighborhood in Evergreen Springs.</i>
What hazards or potential hazards can be identified?	<i>Based on the image metadata and object analysis, the following hazards or potential hazards can be identified: 1. Flooding: The primary hazard is the large-scale flood, which can lead to property damage, infrastructure issues, and potential threats to human safety. 2. Submerged vehicles: The presence of vehicles (cars and land vehicles) in the flooded area indicates that they may be damaged or pose a risk to drivers and passengers. 3. Impassable roads: The flooding may make roads, such as the parking lot mentioned in the MS Azure image description, difficult or impossible to navigate, disrupting transportation and access to services. 4. Damaged property: The presence of water, houses, and trees in the image analysis suggests that residential properties and vegetation may be affected by the flood. 5. Potential electrical hazards: Flooding can lead to electrical hazards if water comes into contact with electrical systems or appliances. Please note that this analysis is based on the available metadata and image analysis, and additional hazards may be present that are not mentioned here.</i>
What is the flood depth when putting the water in relation to different identifiable objects?	<i>Based on the provided metadata and image analysis, we cannot determine the exact flood depth in relation to different identifiable objects. However, it's worth noting that the flood seems to have reached a parking lot, which suggests that the flood has covered ground level areas. Further assessment of the actual image would be required to provide more accurate information about the flood depth in relation to specific objects.</i>
What are the best routes for access and egress?	<i>Given the limited information provided in the metadata and image analysis, it is not possible to suggest the best routes for access and egress in the affected area. To determine the best routes, it would be necessary to have a comprehensive understanding of the local road network, current road conditions, and the extent of the flood. For such purposes, consulting local authorities, emergency services, or real-time mapping applications like Google Maps or Waze would provide more accurate and up-to-date information on the best routes for access and egress.</i>
How many casualties are there and what condition are they in?	<i>Based on the provided metadata and image analysis, there is no information available about casualties or their conditions. To obtain accurate information on casualties and their conditions, it would be necessary to consult local authorities, emergency services, or news reports from the area.</i>
Can you identify people?	<i>Based on the provided metadata and image analysis, there is no mention of people being present in the image. It is not possible to identify any individuals without analyzing the actual image and having information about people present in the scene.</i>
Which and how many emergency responder assets/personnel are required or are already on-scene?	<i>The provided metadata and image analysis do not give any information about emergency responder assets or personnel required or already on-scene. To determine the appropriate number and type of emergency responders needed, it would be necessary to consult local authorities, emergency services, or news reports from the area. The specific requirements would depend on factors such as the extent of the flood, number of affected individuals, and the severity of the situation.</i>

**Table 3: Incident report generated via GPT-4 (AI text in blue italics). Some questions omitted for space.**

was similar (70 cm). However, when we compared the height of the flood to the car tire, the model suggested that the flood was at least 6-8 inches (15-20 cm) deep. This “*at least*” estimator can help set a threshold, i.e., emergency services can activate different response plans above a certain level and deploy teams to the incident.

Based on these findings, we propose the following approach to estimate water depth more accurately: Classifying identifiable vehicles into three main types, such as *small car*, *medium car*, and *SUV*, and assigning average values for tire height. Suppose we can identify the vehicle’s make and model; we can then analyse how high the floodwater is in relation to the tire to provide emergency services with a quantitative indication of severity. Nevertheless, it is crucial to point out that the quantitative measures given are a very rough estimate, and the actual water depth could be significantly different depending on various factors, such as the slope of the road, the tire size, and the perspective of the photo. As a next step, we aim to adopt further image segmentation techniques that can help us to obtain further information about flood depth.

### 3.2 Transferability

We can apply the workflow presented in this study to image data from other climate-related hazard events. For example, we used our model to assess local flood events in Newcastle upon Tyne (UK). For this purpose, publicly available images from Twitter and online newspaper outlets and user image descriptions serve as input data. The analysis results helped us better understand the requirements for input data (such as most suitable image parameters and provided image descriptions) and suggest best practices for future model work that can help generate more accurate output results. Furthermore, the suggested M/ETHANE report template can be enhanced and customized by providing additional questions relevant to other emergency agencies.

### 3.3 Architecture Considerations

Executing machine learning models may be difficult given the data and compute needs; however, wireless networks may also be degraded during a disaster. Since we assume these are the only technically challenging components of our form filling system, there are two classes of system architecture: edge models and cloud models. Since compute may be limited on a mobile device, uploading the media to a cloud service to offload model inference removes this constraint. This allows the use of even larger models than could fit or execute on the device. These are all desirable attributes, with the sole drawback being the reliance on a network. Under network constraints deploying small or compressed models to the cellphone will reduce the reliance on the network.

Battery use is also an important consideration, as power outages may prevent charging, and we do not want to keep a device at a charger during a disaster if a charger can even be found. Running a large model on a smartphone would use battery, as would the device’s radios when pushing large image files through a highly degraded network. It is unclear which of these would use more energy, especially as both may be performed by dedicated hardware which would vary by device type.

Between the two architectures, we believe cloud models are preferable. First, during a full network outage no mobile device

reporting is possible, so we only consider cases of a network being degraded in ways that limit bandwidth. Second, since we assume that machine learning models are larger than their input media, when bandwidth is limited to a level that prevents image uploads the network would also be unable to transfer the machine learning model. In this scenario, the only way to perform camera-first form filling would be to have the models on the device before the disaster impacts network services. This may be possible if an mobile application is deployed, but a web-based system would be unable to do this. Third, we believe that reports containing images are more useful. If either cloud models or image-including reports are used, then both can rely on the same uploaded image.

### 3.4 Limitations

While promising, our proposed workflow may be subject to a variety of limitations, which can inspire future work. For example, our approach requires higher broadband capacity to transmit images, or capable compute resources to execute the machine learning models on-device, depending on the deployed system architecture. Second, while the availability of cellular coverage for smartphones is a critical dependency, we note the emerging trend of smartphone use in such applications [2] and the focus on resilient communication infrastructure as an active area of work [7]. Third, our pipeline relies on the training of task-specific neural networks to perform computer vision tasks. This requires the curation of training datasets which could be hard-to-find, and the possibly high compute requirements of training and fine-tuning a model. Furthermore, AI models can produce misclassifications, erroneous statements, and hallucinations during report generation, which can potentially be detrimental to emergency response. While humans produce erroneous statements too, the patterns of human-produced errors are somewhat well-understood, unlike AI-generated errors.

Our analysis of outputs generated by the AI services used in our pipeline identified some weaknesses that require further investigation in future development cycles: When we asked for the number of cars visible in the image, GPT-3 initially indicated that four cars were visible, although only three cars were detected. When asked about the geocoding of addresses, GPT-3 did not provide an accurate answer compared to the geocoding service used. When determining the floodwater depth estimates, we considered different approaches to estimating the floodwater involving additional information or measurements. We expect these limitations to be addressed both by improvement of the individual components, and also by adding error-checking steps for the model-generated outputs.

## 4 CONCLUSION AND FUTURE WORK

Building shared situational awareness is critical for emergency response agencies to respond effectively to an incident. The current incident reporting process can be lengthy and time-consuming for impacted citizens, impeding emergency response. This paper aims to simplify and accelerate this process to gather higher-quality information and reduce the burden on emergency services. This work demonstrates how we can reduce the friction in generating structured data (in this case, a hazard impact assessment report) using a camera-first approach. As a proof-of-concept, we demonstrate how such data pipelines can be built using off-the-shelf AI

services, including Google's Teachable Machine, Microsoft Azure Cognitive Image API, and OpenAI's GPT models. In the future, we aim to integrate the current analysis output into geospatial solutions and complement camera-extracted information with external data sources, such as real-time data about weather and traffic. Additionally, we would like to evaluate modern multimodal models which combine computer vision and NLP, allowing us to answer more complex questions than possible with object detection alone.

## ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. National Science Foundation under Grant Number 1910356, the United Kingdom Research and Innovation's (UKRI) Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Geospatial Systems under Grant Number EP/S023577/1, and the Ordnance Survey of Great Britain.

## REFERENCES

- [1] Helen Sydney Adams. 2022. *How Tech-Enhanced Call Handling Can Help Ambulance Services*. Retrieved March 7, 2023 from <https://healthcare-digital.com/technology-and-ai/how-tech-enhanced-call-handling-can-help-ambulance-services>
- [2] Daniel J Bachmann, Nathan K Jamison, Andrew Martin, Jose Delgado, and Nicholas E Kman. 2015. Emergency Preparedness and Disaster Response: There's an App for That. *Prehospital and Disaster Medicine* 30, 5 (2015), 486–490.
- [3] Dataminr. 2023. *First Alert*. Retrieved March 7, 2023 from <https://www.dataminr.com/first-alert>
- [4] Ericsson. 2022. *Ericsson Mobility Report. Number of Smartphone Subscriptions Worldwide from 2016 to 2021, with Forecasts from 2022 to 2027*. Retrieved March 17, 2023 from <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/>
- [5] Federal Emergency Management Agency (FEMA). 2023. *National Flood Hazard Layer*. Retrieved March 27, 2023 from <https://hazards-fema.maps.arcgis.com/apps/webappviewer/index.html?id=8b0adb51996444d4879338b5529aa9cd>
- [6] Behzad Golparvar and Ruo-Qian Wang. 2020. -Supported Citizen Science to Monitor High-Tide Flooding in Newport Beach, California. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities*. 66–69.
- [7] Karina Gomez, Leonardo Goratti, Tinku Rasheed, and Laurent Reynaud. 2014. Enabling Disaster-resilient 4G Mobile Communication Networks. *IEEE Communications Magazine* 52, 12 (2014), 66–73.
- [8] Amy/ WUSF Public Media Green. 2022. *Florida Commits \$1 Billion to Climate Resilience, after Hurricane Ian, Some Question the State's Development Practices*. Retrieved March 27, 2023 from <https://wusfnews.wusf.usf.edu/environment/2022-10-13/florida-commits-1-billion-climate-resilience-hurricane-ian-some-question-state-development-practices>
- [9] Hexagon. 2023. *Detect Complex Emergencies Sooner with the Help of Assistive AI*. Retrieved March 7, 2023 from <https://hexagon.com/products/hxgn-oncall-dispatch-smart-advisor>
- [10] JESIP. 2023. *Introduction to the Joint Doctrine*. Retrieved March 27, 2023 from <https://www.jesip.org.uk/joint-doctrine/introduction-to-the-joint-doctrine/>
- [11] JESIP. 2023. *M/ETHANE*. Retrieved March 17, 2023 from <https://www.jesip.org.uk/joint-doctrine/m-ethane/>
- [12] Lilong Jiang, Protiva Rahman, and Arnab Nandi. 2018. Evaluating Interactive Data Systems: Workloads, Metrics, and Guidelines. In *Proceedings of the 2018 International Conference on Management of Data*. ACM, 1637–1644.
- [13] HPCI Media. 2022. *Struggling Ambulance Services Ignoring the Potential of AI to Relieve Pressures*. Retrieved March 7, 2023 from [https://buildingbetterhealthcare.com/news/article\\_page/Struggling\\_ambulance\\_services\\_ignoring\\_the\\_potential\\_of\\_AI\\_to\\_relieve\\_pressures/203185](https://buildingbetterhealthcare.com/news/article_page/Struggling_ambulance_services_ignoring_the_potential_of_AI_to_relieve_pressures/203185)
- [14] Microsoft. 2022. *End-To-End Geospatial Storage, Analysis, and Visualization*. Retrieved April 17, 2023 from <https://learn.microsoft.com/en-us/azure/orbital/geospatial-reference-architecture>
- [15] United Nations. 2015. Transforming our World: the 2030 Agenda for Sustainable Development. Retrieved March 17, 2023 from <https://sdgs.un.org/2030agenda>
- [16] Newzoo. 2022. *Newzoo Global Mobile Market Report. Number of smartphone users by leading countries in 2021 (in millions)\**. Statista. Retrieved March 17, 2023 from <https://www.statista.com/statistics/748053/worldwide-top-countries-smartphone-users/>
- [17] NOAA. 2022. NOAA National Centers for Environmental Information, Monthly National Climate Report for September 2022. <https://www.ncei.noaa.gov/access/>

monitoring/monthly-report/national/202209

- [18] The City of Columbus. 2023. *Service Request Portal*. Retrieved March 17, 2023 from <https://www.columbus.gov/311serviceportal/>
- [19] Department of Homeland Security. 2017. *National Incident Management System. Third Edition/*. Retrieved March 27, 2023 from [https://www.fema.gov/sites/default/files/2020-07/fema\\_nimsDoctrine-2017.pdf](https://www.fema.gov/sites/default/files/2020-07/fema_nimsDoctrine-2017.pdf)
- [20] British Association of Public Safety Communications Officials. 2023. Context is Everything. *BAPCO Journal* 20, 1 (2023), 36–37.
- [21] City of Toronto. 2023. *Create Service Request*. Retrieved March 17, 2023 from <https://www.toronto.ca/home/311-toronto-at-your-service/create-a-service-request/#category>
- [22] Intergovernmental Panel on Climate Change (IPCC). 2022. Climate Change 2022: Impacts, Adaptation and Vulnerability Summary for Policymakers Headline Statements. <https://www.ipcc.ch/report/ar6/wg2/resources/spm-headline-statements/>
- [23] Protiva Rahman and Arnab Nandi. 2019. Transformer: a Database-driven Approach to Generating Forms for Constrained Interaction. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 485–496.
- [24] RedBox. 2023. *AI Powered Speech Analytics Supporting Emergency Services*. Retrieved March 7, 2023 from <https://www.redboxvoice.com/sectors/public-safety-critical-communications/public-safety-analytics/>
- [25] LLB Reporter. 2022. *Latest NHS Figures Show Ambulance Services Face Toughest Winter Yet*. Retrieved March 17, 2023 from <https://londonlovesbusiness.com/latest-nhs-figures-show-ambulance-services-face-toughest-winter-yet/>
- [26] Enrico Rukzio, Chie Noda, Alexander De Luca, John Hamard, and Fatih Coskun. 2008. Automatic Form Filling on Mobile Devices. *Pervasive and Mobile Computing* 4, 2 (April 2008), 161–181. <https://doi.org/10.1016/j.pmcj.2007.09.001>
- [27] Salih Sarp, Murat Kuzlu, Mecit Cetin, Cem Sazara, and Ozgur Guler. 2020. Detecting Floodwater on Roadways from Image Data Using Mask-R-CNN. In *2020 International Conference on INnovations in Intelligent Systems and Applications (INISTA)*. IEEE, 1–6.
- [28] Cem Sazara, Mecit Cetin, and Khan M Iftekharuddin. 2019. Detecting Floodwater on Roadways from Image Data with Handcrafted Features and Deep Transfer Learning. In *2019 IEEE intelligent transportation systems conference (ITSC)*. IEEE, 804–809.
- [29] Farhana Sultana, Abu Sufian, and Paramartha Dutta. 2018. Advancements in Image Classification Using Convolutional Neural Network. In *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 122–129.
- [30] United States Geological Survey (USGS). 2023. *Coastal Change Hazard Portal*. Retrieved March 7, 2023 from <https://marine.usgs.gov/coastalchangehazardsportal/info/>
- [31] United States Geological Survey (USGS). 2023. *National Water Dashboard*. Retrieved March 7, 2023 from <https://dashboard.waterdata.usgs.gov/app/nwd/en/>
- [32] Kristina Wolf, Richard Dawson, Jon Mills, Phil Blythe, Jeremy Morley, and Arnab Nandi. 2023. Supporting Geospatial Climate Hazard Reporting using Computer Vision and Text Generation of Social Media Imagery. 31st Annual Geographical Information Science Research UK Conference (GISRUK), Glasgow, Scotland. <https://doi.org/10.5281/zenodo.7834890>
- [33] Kristina Wolf, Richard J Dawson, Jon P Mills, Phil Blythe, and Jeremy Morley. 2022. Towards a Digital Twin for Supporting Multi-Agency Incident Management in a Smart City. *Scientific reports* 12, 1 (2022), 16221.