# AI for HADR: Progress and Opportunities

*Ross Luo*
*Michael Laielli*
*Giscard Biamby*
*Adam Loeffler*
*Trevor Darrell, Ed.*
*John F. Canny, Ed.*

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2020-233
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2020/EECS-2020-233.html

December 19, 2020

Acknowledgement

First and foremost, I would like to thank Dr. Trevor Darrell for your vision and support the entire way.

To the entire HADR team - Trevor Darrell, Mike Laielli, Giscard Biamby, and Adam Loeffler - we've come a long way since when we first started. None of this would have been possible without you. I've learned so much from you all and I look forward to seeing your future contributions.

To friends, mentors, and my parents, thank you for all the nudges, support, and advice along the way. It has been a long journey to get here but it was not a lonely one.

## AI for HADR: Progress and Opportunities

by Ross Luo

## Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

**Committee:**

Professor Trevor Darrell
Research Advisor

Dec 17, 2020

* * * * * * *

Professor John Canny
Second Reader

Dec 18, 2020

AI for HADR: Progress and Opportunities

by

Ross Luo

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Trevor Darrell, Chair
Professor John Canny

Fall 2020

AI for HADR: Progress and Opportunities

Abstract

AI for HADR: Progress and Opportunities

by

Ross Luo

Master of Science in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Trevor Darrell, Chair

In recent years, the AI research and the humanitarian assistance and disaster response (HADR) communities have sought to collaborate together: There is a growing desire in the AI research community to transition state of the art research towards endeavors for social good. Likewise, the HADR community, comprised mainly of NGO's, governments, and not-for-profit entities, has historically been insulated from the latest technological advances and welcomes infusions of technical insights. Part I of this work describes the origins, progress, and takeaways of these collaborations.

Part II details some of my team's ongoing contributions in this area. Notably, we are the first to adapt structured Gaussian filters to the object detection task. We evaluate our CenterNet-DLA detector with spherical Gaussian filters on COCO and the xView overhead object detection dataset and achieve performance comparable to one with free-form deformable convolution filters while utilizing fewer dynamic parameters.

# Contents

## Acknowledgments

First and foremost, I would like to thank Dr. Trevor Darrell for your vision and support the entire way.

To the entire HADR team - Trevor Darrell, Mike Laielli, Giscard Biamby, and Adam Loeffler - we've come a long way since when we first started. None of this would have been possible without you. I've learned so much from you all and I look forward to seeing your future contributions.

To friends, mentors, and my parents, thank you for all the nudges, support, and advice along the way. It has been a long journey to get here but it was not a lonely one.

# Part I

# AI for HADR

# Chapter 1

# Overview

Part I describes the origins, progress, and takeaways of recent academic initiatives aiming to apply the latest AI advances to solve challenges faced by the HADR community. Many of these challenges are not new but new AI advances have ignited a wave of enthusiasm in tackling them. This work is meant to showcase high level trends and is not meant to be exhaustive. Previous initiatives targeting these challenge areas are outside the scope of this work.

# Chapter 2

# Introduction

Research advances in the past decade have generated optimism about AI applications. CNN's now outperform humans [13] in the ImageNet classification task; these and other computer vision advances have kick-started industries aimed at automating aspects of human perception. Likewise, speech recognition algorithms now match or surpass human performance on standard benchmarks [47] and have been integrated into the user interfaces of everything from personal phones to smart homes to automobiles. Reinforcement learning advances have enabled AI to attain or exceed human abilities in a wide range of tasks that previously were not thought to be possible such as complex games [6, 43]. As a result of these developments, market analysts project that the emerging AI industry will contribute between 6-16 trillion dollars to the global economy in the next 5-15 years [12, 32, 33].

On the flip side, there is growing concern about the negative consequences of these advances. Generative networks are now able to synthesize "deep fakes", realistic text, images, videos, and audio that could be utilized for misinformation and criminal activities [1, 26, 29]. Parallel developments in AI algorithms, compute, storage, and cloud infrastructure have enabled the engineering of mass surveillance systems. Algorithms operating on social media platforms optimizing for engagement have been blamed for political polarization and echo chambers. There is concern that many AI advances are dual-use, able to be adapted for civilian as well as military purposes, and are contributing to an AI arms race [39]. Finally, the themes of AI fairness, interpretability, and explainability are now major topics of discussion in industry, academia, governments, and NGO's as algorithms are increasingly being allowed to make consequential and lasting decisions on people's lives [31].

These negative consequences have led to calls in the AI community to make sure more is done to so that AI advances lead to positive societal outcomes. As early as 2017, industry, academia, governments, and NGO's have come together and begun pursuing initiatives under banners such as "AI for Good" [3], "AI for Social Good" [17, 19, 46], "Computer Vision for Global Challenges" [8], and most recently "AI for HADR" [2].

## 2.1 What is HADR

Disasters, natural and man-made, strike countries both rich and poor. The stakeholders that respond to these calamities, providing food, supplies, monetary aid, manpower, leadership, organization, technical expertise, security, and much more, collectively represent the humanitarian assistance and disaster response (HADR) community.

HADR operations can be directed domestically or abroad. In contrast to developmental aid which aims to build long-term resilience to mitigate potential disasters, HADR encompasses the short-term steps taken immediately before, during, and after a disaster, encompassing evacuation, search and rescue, aid, and recovery. HADR overlaps with humanitarian aid, disaster recovery, disaster relief, and disaster management.

The HADR community is comprised of government entities, commercial interests, not-for-profit organizations, NGO's such as the Red Cross [9], and intergovernmental entities such as the UN [27]. In the United States, HADR governmental entities include domestic-facing stakeholders such as first responders, decision makers at various levels of government, FEMA, and the National Guard as well as foreign-facing stakeholders such as USAID,the Department of State, and the Department of Defense.

## 2.2 AI for HADR

In 2019, NeurIPS was host to the first AI for HADR workshop. The purpose of the workshop was to "establish meaningful dialogue between the AI and HADR communities" and to help transition "the research created by the NeurIPS community to real world humanitarian issues" [2]. Organizers and participants of the workshop included AI researchers from academia and industry on the one hand and practitioners from not-for-profits, NGO's and governmental entities on the other. The workshop showcased ongoing collaborations in areas such as flood and fire prediction, damage assessment, and social media analysis during disasters. A subsequent workshop was also hosted at NeurIPS 2020.

## 2.3 Related Initiatives

Prior to the formation of the AI for HADR workshop, the AI for Social Good and Computer Vision for Global Challenge workshops were forums for AI for HADR work.

### AI for Social Good

In 2018, NeurIPS hosted its first AI for Social Good workshop. The organizers framed the workshop's focus to be "on social problems for which artificial intelligence has the potential to offer meaningful solutions" and solicited papers targeting themes inspired by the UN Sustainable Development Goals (SDGs). These themes were: education, protecting democracy, urban planning, assistive technology for people with disabilities, health, agriculture, environmental sustainability, economic, social, and gender inequality, social welfare and justice, ethics, privacy, and security. The workshop brought together researchers, practitioners, and philanthropists together into the same room to to build a community that could share their successes and failures, identify evaluation criteria, build new tools and datasets, and inspire others to join the mission [4]. Subsequent AI for Social Good workshops were organized for ICLR 2019, ICML 2019, and NeurIPS 2019.

## Computer Vision for Global Challenges (CV4GC)

The inaugural and only Computer Vision for Global Challenges (CV4GC) workshop was hosted at CVPR 2019. Like AI for Social Good, CV4GC targeted priorities like the UN SDG's but narrowed its focus to computer vision applications [8].

# Chapter 3

# Review of Recent Progress

## 3.1 Overall Trends

The scope of this review is limited to HADR papers accepted at the AI for Social Good, Computer Vision for Global Challenges, and AI for HADR workshops.To formally extract insights about the scope of the efforts featured at these workshops, text mining was performed on the workshop websites. Titles of accepted papers, names of authors, and author institutions were aggregated, cleaned, and analyzed.

To begin, Table 3.1 tallies the number of accepted HADR papers at these workshops. It is notable that 2019 was the height for the number of accepted papers. NeurIPS 2020 did not host an AI for Social Good workshop but it did host an AI for HADR workshop. The reason for all of this is unclear, but a center of focus for researchers in 2020 has been topics related to COVID-19.

| Number of Accepted HADR Papers by Workshop | | | |
|---|---|---|---|
| **Workshop** | **Conference** | ***Accepted HADR Papers** | **Total Accepted** |
| AI for Social Good | NeurIPS 2018 | 9 | 40 |
| | ICLR 2019 | 2 | 24 |
| | ICML 2019 | 2 | 33 |
| | NeurIPS 2019 | 2 | 41 |
| CV for Global Challenges | CVPR 2019 | 5 | 14 |
| AI for HADR | NeurIPS 2019 | 13 | |
| | NeurIPS 2020 | 11 | |

Table 3.1: A tally of HADR papers. *HADR papers were qualified based on whether abstracts mentioned applications that fall into areas covered by the HADR workshops. For list of works categorized as HADR, see Appendix A.

To gain insight into where AI for HADR research has been concentrated, author affiliations were tallied. Table 3.2 ranks the top author affiliations. Notably, industry and academia are both invested in AI for HADR research.

| Top Author Affiliations in HADR Papers | | |
|---|---|---|
| **Author Affiliation** | **Type of Institution** | ***Number of Appearances** |
| Google | Industry | 45 |
| Carnegie Mellon University | Academia | 16 |
| CrowdAI | Industry | 7 |
| Mila | Academia | 6 |
| Facebook | Industry | 6 |

Table 3.2: Top author affiliation appearances in HADR papers. *Tallies are per author per paper. Repeat authors and papers with a large number of authors inflate the tallies.

## 3.2 Common Themes

Because the scope of the three workshops have significant overlap, several initiatives spanned multiple workshops and many authors appeared in multiple workshops. Notably, a flood prediction project spanned four papers across three workshops.

To gain insights into the contents of the HADR papers, abstracts were used. We first use the abstracts to tag each paper with an application area, where appropriate. Table 3.3 lists top 5 common application areas. Most of the damage assessment, mapping, and wildfire prediction papers leveraged computer vision methods. NLP was leveraged in works that aimed to mine social media in order to provide early disaster warning.

| **Top Application Areas** | |
|---|---|
| **Application Area** | **Number of Papers** |
| Damage Assessment | 10 |
| Mapping | 7 |
| Flood Modeling | 6 |
| Wildfire Prediction | 5 |
| Social Media Analysis | 4 |

Table 3.3: Top application areas targeted by accepted papers

The contents of the abstracts were then mined. Table 3.4 is a filtered list of the most frequent words and phrases found in the abstracts.

| **Most Frequent Words and Phrases** | |
|---|---|
| **Word or Phrase** | ***Number of Appearances** |
| Satellite | 30 |
| Damage | 28 |
| Building | 20 |
| Flood | 19 |
| Fire | 14 |
| Crowd | 8 |
| Social Media | 7 |

Table 3.4: Filtered list of most frequent words and phrases appearing in abstracts. *Tally refers to total number of appearances in abstracts. Occurrences in the same abstract are double counted.

## 3.3 Common Challenges

There are three common challenges that have been faced by the AI for HADR community: 1. The Researcher-Practitioner Divide. 2. Dataset Availability 3. Short-Term Collaborations Without Sustained Impact.

### Researcher-Practitioner Divide

All three workshops operated with similar formats, featuring a mix of poster presentations, oral presentations, invited talks, and speaker panels. Ostensibly, the rationale for this is captured best by the stated motivation for the CV4GC workshop [8]:

> "We argue that one of the main obstacles is the disconnection between domain experts: those who are close to the problems on the ground, and those who have knowledge about technical solutions. This disconnect might be driven by geographical divide, differences in language and taxonomy, or might come from the lack of a accessible forum to find each other. We propose this initiative as a first step to bridge the gap between these two communities."

AI for HADR adopts a similar purpose [2]:

> "We intend to establish meaningful dialogue between the Artificial Intelligence (AI) and Humanitarian Assistance and Disaster Response (HADR) communities. By the end of the workshop, the NeurIPS research community can learn the practical challenges of aiding those in crisis, while the HADR community can get to know the state of art and practice in AI"

The structures of the workshops were designed to mitigate long-standing divides and to build a community with meaningful dialogue. Invited talks were opportunities for practitioners to describe problem statements and provide datasets and resources to translate research insights into practical outcomes. Poster and oral presentations were opportunities for researchers to share lessons learned.

### Dataset Availability

Dataset availability is a prerequisite for any meaningful machine learning work. All of the top application areas highlighted in Table 3.3 were enabled by dataset collaborations between researchers and practitioners.

In flood prediction for example, Google recognized the scarcity of publicly available data [28] and collaborated with India and Bangladesh to access real-time and historical water level measurements, leading to a long-term project to provide flood warnings.

In damage assessment, the xBD dataset [18] was a collaboration between CrowdAI, a commercial company, CMU's SEI, and the Department of Defense. It was unveiled at the first AI for HADR workshop at NeurIPS 2019 and subsequently enabled multiple building damage assessment submissions to the AI for HADR workshop at NeurIPS 2020.

### Short-Term Collaborations Without Sustained Impact

AI for HADR branched off from the AI for Social Good movement but shares some of the same hurdles. As Tomasev et al. note in their review of the AI for Social Good movement, short-term initiatives like workshops are critical in "gathering momentum and bringing application-domain experts together with AI researchers" but they recommend establishing "long-term collaborations between application-domain experts and AI researchers and form[ing] deep integrated partnerships that allow for enough time to reach good practical solutions" [46].

One of the most important prerequisites for long-term collaborations is incentive-alignment. Some of the high profile projects in AI for HADR, such as Google's flood prediction initiatives and Cal Fire's initiatives with CrowdAI and the California ANG [14], have evolved into deep partnerships because stakeholders on all sides have managed to find common incentive. In recent years, both floods and wildfires have have inflicted significant material and human cost. In particular, the 2018 California wildfire season was the worst in the state's recorded history and was also the worst in the nation for that year [38]. India's 2019 monsoon season recorded the most monsoonal rainfall in the last 25 years and led to the injury and displacement of more than 2.5 million people [30]. These natural disaster events forced governmental entities to invest in mitigation measures and offered AI researchers opportunities to apply state of the art research to effect positive outcomes.

## 3.4 Conclusions

AI for HADR branched off from the AI for Social Good movement and has led to some meaningful collaborations between researchers and practitioners in areas such as flood and wildfire prediction and damage assessment. As noted in Tomasev et al. note, the trend of applying tech to social domains is not new [46]. What is new with the new AI-centric movements is that datasets have now become a prerequisite. The most successful collaborations in the AI for HADR space have been enabled by first meaningfully connecting researchers with practitioners, making sure researchers have access to the necessary datasets, and then making sure incentives are aligned for long-term collaboration and dialogue.

# Part II

# Summary of Contributions

# Chapter 4

# Overview

Part II is a summary of some the team's contributions to the AI for HADR community. Our aim was to advance state of the art performance on overhead imagery datasets by evaluating new methods on the xView dataset.

With this goal in mind, we adapted two scale-adaptive object detection methods, deformable convolutions and spherical Gaussian filters, to a CenterNet-DLA detector and benchmarked them against a SSD baseline. Our hope is that insights gained from our work can be adapted to improve disaster response capabilities.

# Chapter 5

# xView Dataset

The xView dataset was released in 2018 as part of the xView detection challenge with the aim of advancing solutions for national security and disaster response. It is an overhead object detection dataset with image chips spanning more than 1400 $km^2$ of Earth's surface and labeled with roughly 1 million objects across 60 classes. Image chips were sourced from Digital Globe WorldView-3 satellites with 0.3m ground sampling distance. The raw dataset chips each cover roughly 1 $km^2$ in ground area and measures roughly $3000^2$ pixels. Figure 5.1 shows some example image chips.

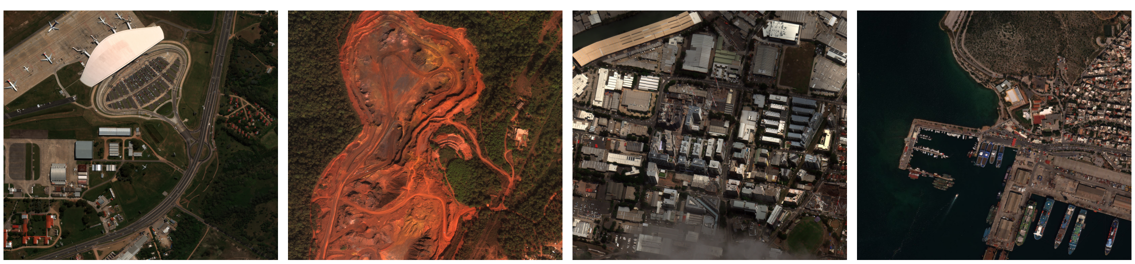

Figure 5.1: Example Xview image chips showcasing the geographic span of the dataset

## 5.1 Dataset Characteristics

The xView dataset [22] possesses many characteristics that are common to overhead object detection datasets. These include the presence of dense labels, object occlusions, small objects, and imbalances in label density, detection scale, and class counts [42]. These characteristics make xView a great proxy for other overhead datasets but also make it a challenging dataset to tackle.

Some of these characteristics are the result of the data curation strategy. xView image chips were sourced to minimize bias across areas of interest (AOIs) [22]. Example AOI's include mines, ports, airfields, and coastal, inland, urban, and rural regions. However, due to the emphasis on minimizing AOI bias, other data imbalances were created. Class counts are extremely imbalanced, with the "building" and "small car" labels dominating instance counts due to their prevalence in dense urban AOI's. See figure 4.2 for class instance count distribution.

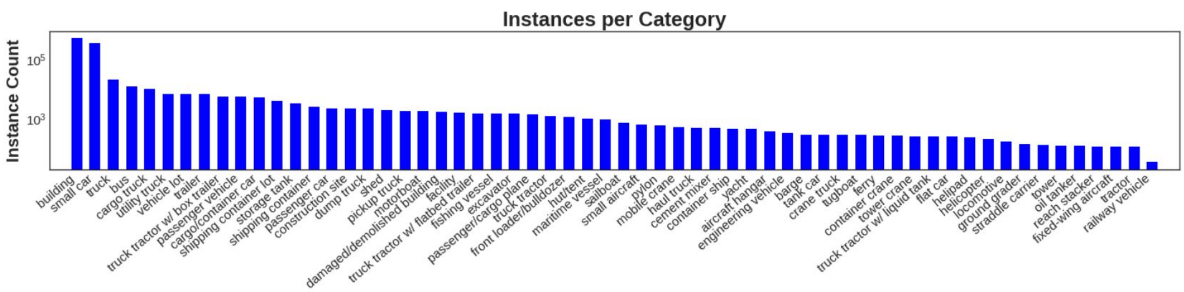


Figure 5.2: Figure from the xView paper [22] documenting class instance count distribution

Other characteristics are a function of the inherent constraints of satellite imagery. The WorldView-3 satellite produces some of the most high resolution commercial imagery available with a ground sampling distance of 0.3m per pixel (the federal limit is 0.25m). Even then, objects such as cars and vehicles still end up being sometimes only a few pixels across. This makes the classification task difficult even for human eyes. The resolution constraint also means that image chips span wide swaths of area, leading to urban image chips containing thousands of overlapping labels coexisting with rural image chips with no labels.

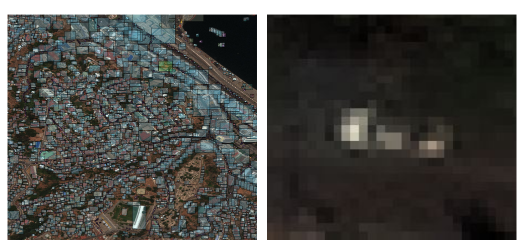

Figure 5.3: Example of dense detections in an urban chip and the low resulting resolution of a small object (car)

All of the above characteristics limit the ability for object detection algorithms to generalize effectively. As a result, even the 1st place winner of the xView challenge only achieved a mAP of 0.28 [40]. Compare this to literature state of the art performance on the MS COCO object detection dataset, which currently lies at 0.55 [45].

# Chapter 6

# Object Detection Efforts

## 6.1 Overview

To tackle xView's scale variance, we adapted a CenterNet detector with a DLA-34 backbone and compared the performance of free-form deformable convolution filters to spherical Gaussian filters. We are the first to apply structured Gaussian filters to a detection task. The two yielded nearly identical mAPs in the xView and COCO datasets but Gaussian filters utilize fewer dynamic parameters. Both of these methods outperformed a baseline SSD-512 detector by 0.02 mAP.

## 6.2 Introduction

Prior to the proliferation of CNN's in object detection, state of the art detectors utilized sliding window and selective search methods for localizing objects within an image [5, 16] and handcrafted feature extractors with HOG [11] and SIFT [25] to help classify them.

After AlexNet won the ImageNet classification challenge in 2012 [21], RCNN [16] was among the first to demonstrate the advantage of using CNN backbones as feature extractors. CNN backbones have advanced a lot since then. Skip connections (ResNet) [20] and inception modules (GoogLeNet) [44] have allowed for the training of deeper and more accurate networks.

For object localization, RCNN and later Fast RCNN used a selective search based region proposal framework [7] before passing the proposals to a CNN for classification [15, 16]. Faster RCNN replaced selective search with a region proposal network (RPN) that generates proposals based on anchor points [37]. Subsequent single-stage detectors such as SSD [24] and YOLO [34, 35, 36] implemented localization with concepts similar to anchor points. CenterNet replaces anchor points that are exhaustively generated with center points that are placed at the peaks of class heatmaps generated from image features. Bounding boxes are then regressed from the features at the peaks [48]. This is more computational efficient and allows CenterNet to achieve strong speed-accuracy trade off.

To account for the scale and transformation variance of objects and the inherent hierarchical attributes of images, various approaches have been proposed. Historically, image feature pyramids had been used to incorporate image features at different scales. SSD, FPN, and DLA use similar principles to sample and incorporate features from different layers [23, 24, 49]. DLA in particular utilizes hierarchical skip connections. In addition, various proposals have been made to modify the convolution operations in these networks. Normally, convolutional layers are calculated by passing a fixed convolution filter of size k x k across a feature map with some stride. This means that regardless of the task a given layer's features are sampled at fixed locations. Deformable convolutions have been shown to improve accuracy in object detection and segmentation tasks by dynamically altering filter offset based on input features [10]. Structured adaptive receptive fields such as gaussian filters achieve similar performance as the more computational-intensive free-form deformable convolutions but with fewer dynamic parameters [41].

## 6.3 CenterNet Detector

The CenterNet detector treats object detection as a keypoint estimation problem. A heatmap for each class is generated for each image. Peaks of the heatmaps are used to propose center points and the features at those points are used to regress bounding boxes. [48]

### DLA-34 Backbone

We use a DLA-34 backbone for feature extraction. Deep Layer Aggregation (DLA) emphasizes deeper connectivity between layers by adding hierarchical skip connections that incorporate earlier layers more. The result is improved performance, parameter count, and memory usage over baseline models that use traditional skip connections. The approach is similar to that of FCN skip connections and FPN top-down connections. DLA-34 is specifically a ResNet-34 backbone augmented with iterative (IDA) and hierarchical deep aggregation (HDA) blocks (See figure 6.1) [49].

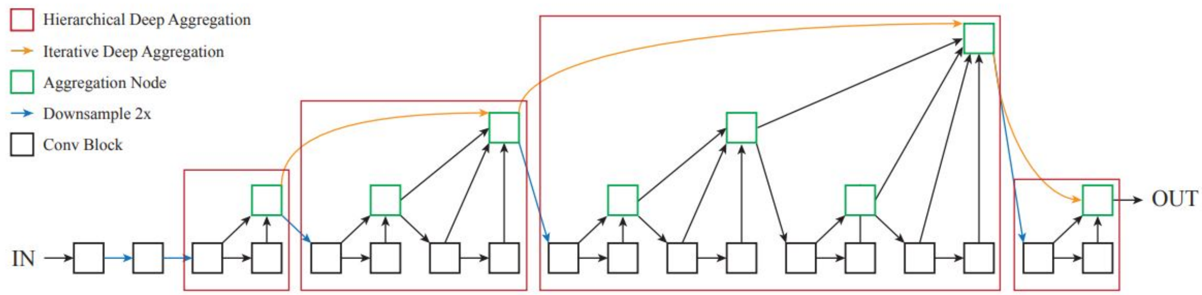


Figure 6.1: Modifications to a ResNet backbone featuring both IDA and HDA blocks. Diagram is from DLA paper [49]

## Deformable Convolutions

Deformable convolutions allow for dynamic adjustments to the receptive fields of convolutional layers at inference time in order to account for scale variations that are inherent in images. This is done by adding 2D offsets to the grid sampling locations of a standard convolution filter. (See figure 6.2)

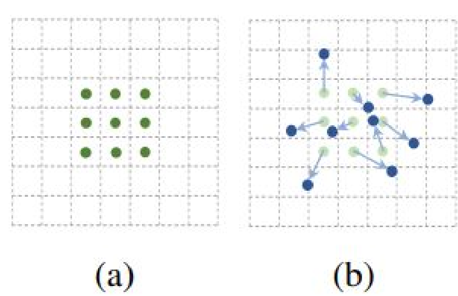


Figure 6.2: A deformable convolution yields a free-form receptive field (b) compared to a fixed receptive field from a traditional convolution (a). Figure is from the DCN paper [10]

These offsets are obtained by passing the same input feature map through an additional convolutional filter to obtain an offset field. The result is a free-form receptive field at each location in a feature map (see Figure 6.3).

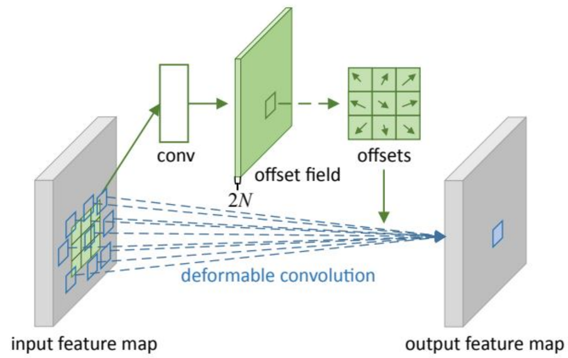


Figure 2: Illustration of 3 × 3 deformable convolution.

Figure 6.3: A deformable convolution involves learning an offset field that is utilized to predict offsets specific to each input. Figure is from the DCN paper [10]

Our DLA-34-Deformable backbone is a DLA-34 backbone with additional skip connections and its 3 x 3 convolutions in the upsampling stages substituted with deformable convolutions (see figure 6.4)

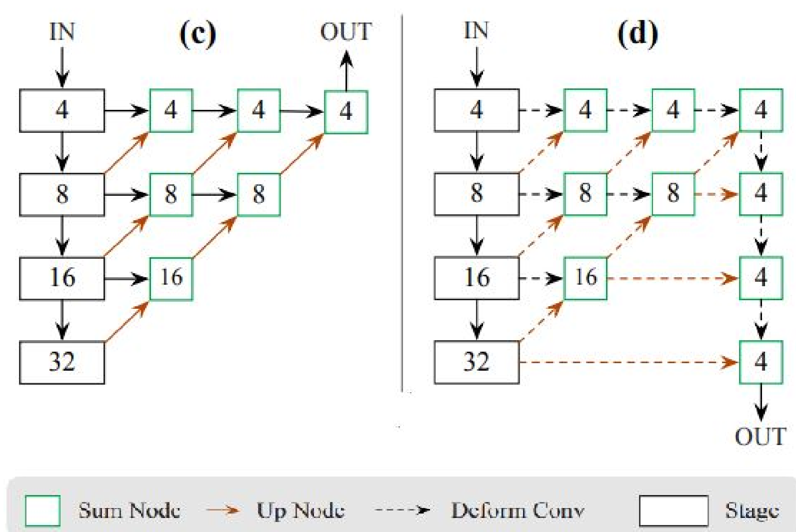


Figure 6.4: DLA-34-Deformable is a DLA-34 with extra skip connections and static convolutions replaced with deformable convolutions. Figure is from the CenterNet paper [48]

### Gaussian Sigma Filter

In contrast to free-form deformable convolution filters, structured Gaussian filters have been proposed to allow for dynamic receptive fields while reducing the number of dynamic parameters [41].

Here, we adopt spherical Gaussian filters where a scaled offset can be specified with a single $\sigma$ parameter through it's covariance $\Sigma$ where $\Sigma = \left(\begin{smallmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{smallmatrix}\right)$

$$G(\mathbf{x};\mathbf{\Sigma}) = \frac{1}{2\pi\sqrt{\det \mathbf{\Sigma}}} e^{-\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x}/2}$$

Figure 6.5 illustrates what a spherical Gaussian filter looks like.

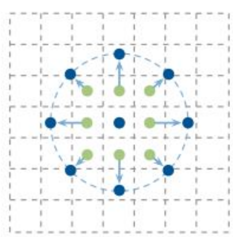

Figure 6.5: Illustration of a spherical Gaussian filter. A single $\sigma$ parameter determines offset scaling through its covariance $\Sigma$

Our DLA-34-Sigma backbone is a DLA-34-Deformable backbone with deformable convolutions replaced with these spherical Gaussian filters. In doing so, we reduce the number of dynamic parameters per convolutional layer from $2k^2$ to 1.

## 6.4 Results and Conclusion

For our experiments, we chose the COCO dataset and SSD-512 detector as baselines. We chipped xView to 512x512 image chips to match the input resolution of COCO.

We trained each detector (SSD-512, CenterNet-DLA-Deformable, CenterNet-DLA-Sigma) to convergence on COCO and xView and collect evaluation results in Table 6.1.

| **Evaluation Results** | | |
|---|---|---|
| **Object Detector** | **COCO mAP** | **xView mAP** |
| SSD-512 | 0.465 | 0.178 |
| CenterNet-DLA-Deformable | 0.485 | 0.198 |
| CenterNet-DLA-Sigma | 0.485 | 0.187 |

Table 6.1: Evaluation results on COCO and xView.

CenterNet DLA outperformed SSD in both COCO and xView. Furthermore, the spherical Gaussian filters were shown to provide most if not all of the benefits of the deformable convolutions while utilizing fewer dynamic parameters ($2k^2$ vs 1 for each convolutional layer).

## 6.5 Future Work

While we show the advantage of CenterNet-DLA modified with adaptive receptive fields over a baseline SSD, additional ablation studies involving CenterNet-DLA without dynamic receptive fields and CenterNet with different backbones could be conducted to determine the relative contributions of the CenterNet detector and the DLA backbone to our performance. To explore performance differences, training times between SSD and CenterNet could be compared. The speed-accuracy trade-off at test time and memory footprint differences for dynamic receptive fields could also be explored.

# Appendix A

# List of Papers Categorized As HADR

*Excludes papers accepted to the AI for HADR workshops

| **Workshop** | **Conference** | **Title** | **Authors** |
| --- | --- | --- | --- |
| AI4SG | NeurIPS18 | A Longitudinal Evaluation of a Deployed Predictive Model of Fire Risk | Jessica Lee et al. |
| AI4SG | NeurIPS18 | ML for Flood Forecasting at Scale | Sella Nevo et al. |
| AI4SG | NeurIPS18 | Intelligent Drone Swarm for Search and Rescue Operations at Sea | Vincenzo Lomonaco et al. |
| AI4SG | NeurIPS18 | From Satellite Imagery to Disaster Insights | Jigar Doshi et al. |
| AI4SG | NeurIPS18 | A Complementary Approach to Improve Wild Fire Prediction Systems | Sriram Ganapathi Subramanian et al. |
| AI4SG | NeurIPS18 | Witnessing atrocities: quantifying villages destruction in Darfur with crowdsourcing and transfer learning | Julien Cornebise et al. |
| AI4SG | NeurIPS18 | Rapid Computer Vision-aided Disaster Response via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery | Tim G. J. Rudner et al. |
| AI4SG | NeurIPS18 | Foundational mapping of Uganda to assist American Red Cross disaster response to floods and pandemics | Alexei Bastidas et al. |

| | | | |
|---|---|---|---|
| AI4SG | NeurIPS18 | Towards Global Remote Discharge Estimation: Using the Few to Estimate The Many | Yotam Gigi et al. |
| AI4SG | ICLR19 | A pipeline for emergency response | Ayan Mukhopadhyay et al. |
| AI4SG | ICLR19 | Disaster Insurance - New parametric contracts based on satellite images | Eric Bouyé et al. |
| CV4GC | CVPR19 | Building High Resolution Maps for Humanitarian Aid and Development with Weakly- and Semi-Supervised Learning | Derrick Bonafilia et al. |
| CV4GC | CVPR19 | Creating xBD: A Dataset for Assessing Building Damage from Satellite Imagery | Ritwik Gupta et al. |
| CV4GC | CVPR19 | DisplaceNet: Recognising Displaced People from Images by Exploiting Dominance Level | Grigorios Kalliatakis et al. |
| CV4GC | CVPR19 | Deep Landscape Features for Improving Vector-borne Disease Prediction | Nabeel Abdur Rehman et al. |
| CV4GC | CVPR19 | Detecting Roads from Satellite Imagery in the Developing World | Yoni Nachmany et al. |
| AI4SG | ICML19 | Crisis Sub-Events on Social Media: A Case Study of Wildfires | Shan Jiang et al. |
| AI4SG | ICML19 | Addressing Novel Sources of Bias for Change Detection on Large Social Networks | Gabriel Cadamuro et al. |
| AI4SG | NeurIPS19 | Large-Scale Landslides Detection from Satellite Images with Incomplete Labels | Masanari Kimura et al. |
| AI4SG | NeurIPS19 | Using News Articles to Model Hepatitis A Outbreaks: A Case Study in California and Kentucky | Marie Charpignon et al. |

# Bibliography


[1] Agarwal, S. et al. "Protecting World Leaders Against Deep Fakes". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. Long Beach, CA: IEEE, June 2019, p. 8. URL: http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.pdf.

[2] *AI + HADR 2020*. URL: https://www.hadr.ai/home.

[3] *AI for Good Global Summit*. URL: https://aiforgood.itu.int/.

[4] *AI for Social Good NeurIPS2018 Workshop*. URL: https://aiforsocialgood.github.io/2018.

[5] Azayev, T. "Object detection in high resolution satellite images". In: (2016). URL: https://dspace.cvut.cz/bitstream/handle/10467/65237/F3-BP-2016-Azayev-Teymur-TEYMUR_AZAYEV_05_2016_KOS.pdf.

[6] Berner, C. et al. "Dota 2 with Large Scale Deep Reinforcement Learning". In: *ArXiv* abs/1912.06680 (2019).

[7] Chunhui Gu et al. "Recognition using regions". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009, pp. 1030–1037. DOI: 10.1109/CVPR.2009.5206727.

[8] *Computer Vision for Global Challenges*. URL: https://www.cv4gc.org/.

[9] Cross, A. R. "Disaster Relief". In: (). URL: https://www.redcross.org/about-us/our-work/disaster-relief.html.

[10] Dai, J. et al. "Deformable Convolutional Networks". In: *CoRR* abs/1703.06211 (2017). arXiv: 1703.06211. URL: http://arxiv.org/abs/1703.06211.

[11] Dalal, N. and Triggs, B. "Histograms of Oriented Gradients for Human Detection". In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* 1 (2005), pp. 886–893. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1467360.

[12] Deloitte. "Global Artificial Intelligence Industry Whitepaper". In: (2019). URL: https://www2.deloitte.com/cn/en/pages/technology-media-and-telecommunications/articles/global-ai-development-white-paper.html.

[13] Dodge, S. F. and Karam, L. J. "A Study and Comparison of Human and Deep Learning Recognition Performance Under Visual Distortions". In: *CoRR* abs/1705.02498 (2017). arXiv: 1705.02498. URL: http://arxiv.org/abs/1705.02498.

[14] Doshi, J. et al. "FireNet: Real-time Segmentation of Fire Perimeter from Aerial Video". In: *CoRR* abs/1910.06407 (2019). arXiv: 1910.06407. URL: http://arxiv.org/abs/1910.06407.

[15] Girshick, R. B. "Fast R-CNN". In: *CoRR* abs/1504.08083 (2015). arXiv: 1504.08083. URL: http://arxiv.org/abs/1504.08083.

[16] Girshick, R. B. et al. "Rich feature hierarchies for accurate object detection and semantic segmentation". In: *CoRR* abs/1311.2524 (2013). arXiv: 1311.2524. URL: http://arxiv.org/abs/1311.2524.

[17] Google. "AI for Social Good". In: (). URL: https://ai.google/social-good/.

[18] Gupta, R. et al. "xBD: A Dataset for Assessing Building Damage from Satellite Imagery". In: *CoRR* abs/1911.09296 (2019). arXiv: 1911.09296. URL: http://arxiv.org/abs/1911.09296.

[19] Hager, G. D. et al. "Artificial Intelligence for Social Good". In: (2017). URL: https://cra.org/ccc/wp-content/uploads/sites/2/2016/04/AI-for-Social-Good-Workshop-Report.pdf.

[20] He, K. et al. "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385 (2015). arXiv: 1512.03385. URL: http://arxiv.org/abs/1512.03385.

[21] Krizhevsky, A., Sutskever, I., and Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by Pereira, F. et al. Vol. 25. Curran Associates, Inc., 2012, pp. 1097–1105. URL: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b
Paper.pdf.

[22] Lam, D. et al. "xView: Objects in Context in Overhead Imagery". In: *CoRR* abs/1802.07856 (2018). arXiv: 1802.07856. URL: http://arxiv.org/abs/1802.07856.

[23] Lin, T. et al. "Feature Pyramid Networks for Object Detection". In: *CoRR* abs/1612.03144 (2016). arXiv: 1612.03144. URL: http://arxiv.org/abs/1612.03144.

[24] Liu, W. et al. "SSD: Single Shot MultiBox Detector". In: *CoRR* abs/1512.02325 (2015). arXiv: 1512.02325. URL: http://arxiv.org/abs/1512.02325.

[25] Lowe, D. G. "Object Recognition from Local Scale-Invariant Features". In: *Proc. of the International Conference on Computer Vision ICCV, Corfu*. 1999.

[26] Mirsky, Y. and Lee, W. *The Creation and Detection of Deepfakes: A Survey*. 2020. arXiv: 2004.11138 [cs.CV].

[27] Nations, U. "Deliver Humanitarian Aid". In: (). URL: https://www.un.org/en/sections/what-we-do/deliver-humanitarian-aid/.

[28] Nevo, S. et al. *ML-based Flood Forecasting: Advances in Scale, Accuracy and Reach.* 2020. arXiv: 2012.00671 [physics.ao-ph].

[29] Nguyen, T. T. et al. *Deep Learning for Deepfakes Creation and Detection: A Survey.* 2019. arXiv: 1909.11573 [cs.CV].

[30] Patel, K. "Unusual Monsoon Season Causes Flooding in India". In: (). URL: https://earthobservatory.nasa.gov/images/145703/unusual-monsoon-season-causes-flooding-in-india.

[31] Perrault, R. et al. "The AI Index 2019 Annual Report". In: *AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA* (2019).

[32] Purdy, M. and Daugherty, P. "How AI Boosts Industry Profits and Innovation". In: (2017). URL: https://www.accenture.com/_acnmedia/Accenture/next-gen-5/insight-ai-industry-growth/pdf/Accenture-AI-Industry-Growth-Full-Report.pdf?la=en.

[33] Rao, A. S. and Verweij, G. "Sizing the prize: What's the real value of AI for your business and how can you capitalise?" In: (2017). URL: https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf.

[34] Redmon, J. and Farhadi, A. "YOLO9000: Better, Faster, Stronger". In: *CoRR* abs/1612.08242 (2016). arXiv: 1612.08242. URL: http://arxiv.org/abs/1612.08242.

[35] Redmon, J. and Farhadi, A. "YOLOv3: An Incremental Improvement". In: *CoRR* abs/1804.02767 (2018). arXiv: 1804.02767. URL: http://arxiv.org/abs/1804.02767.

[36] Redmon, J. et al. "You Only Look Once: Unified, Real-Time Object Detection". In: *CoRR* abs/1506.02640 (2015). arXiv: 1506.02640. URL: http://arxiv.org/abs/1506.02640.

[37] Ren, S. et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". In: *CoRR* abs/1506.01497 (2015). arXiv: 1506.01497. URL: http://arxiv.org/abs/1506.01497.

[38] Romero, D. "California had nation's worst fire season in 2018". In: (). URL: https://www.nbcnews.com/news/us-news/california-had-nation-s-worst-fire-season-2018-n981431.

[39] Russell, S. *Human Compatible: Artificial Intelligence and the Problem of Control.* Penguin Publishing Group, 2019. ISBN: 9780525558620. URL: https://books.google.com/books?id=M1eFDwAAQBAJ.

[40] Sergievskiy, N. and Ponamarev, A. "Reduced Focal Loss: 1st Place Solution to xView object detection in Satellite Imagery". In: *CoRR* abs/1903.01347 (2019). arXiv: 1903.01347. URL: http://arxiv.org/abs/1903.01347.

[41] Shelhamer, E., Wang, D., and Darrell, T. "Blurring the Line Between Structure and Learning to Optimize and Adapt Receptive Fields". In: *CoRR* abs/1904.11487 (2019). arXiv: 1904.11487. URL: http://arxiv.org/abs/1904.11487.

[42] Shermeyer, J. et al. *SpaceNet 6: Multi-Sensor All Weather Mapping Dataset*. 2020. arXiv: 2004.06500 [eess.IV].

[43] Silver, D. et al. "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm". In: *CoRR* abs/1712.01815 (2017). arXiv: 1712.01815. URL: http://arxiv.org/abs/1712.01815.

[44] Szegedy, C. et al. "Going Deeper with Convolutions". In: *CoRR* abs/1409.4842 (2014). arXiv: 1409.4842. URL: http://arxiv.org/abs/1409.4842.

[45] Tan, M., Pang, R., and Le, Q. V. "EfficientDet: Scalable and Efficient Object Detection". In: *CoRR* abs/1911.09070 (2019). arXiv: 1911.09070. URL: http://arxiv.org/abs/1911.09070.

[46] Tomašev, N. et al. "AI for social good: unlocking the opportunity for positive impact". In: *Nature Communications* (2020). URL: https://doi.org/10.1038/s41467-020-15871-z.

[47] Xiong, W. et al. "Achieving Human Parity in Conversational Speech Recognition". In: *CoRR* abs/1610.05256 (2016). arXiv: 1610.05256. URL: http://arxiv.org/abs/1610.05256.

[48] Yu, F., Wang, D., and Darrell, T. "Deep Layer Aggregation". In: *CoRR* abs/1707.06484 (2017). arXiv: 1707.06484. URL: http://arxiv.org/abs/1707.06484.

[49] Yu, F., Wang, D., and Darrell, T. "Deep Layer Aggregation". In: *CoRR* abs/1707.06484 (2017). arXiv: 1707.06484. URL: http://arxiv.org/abs/1707.06484.