



HITACHI
Inspire the Next

災害時の映像解析AI技術ご紹介

映像解析の国際ワークショップTRECVID2020*の
災害映像認識タスクで、トップレベル精度を達成

2021年11月
株式会社 日立製作所

* 主催：アメリカ国立標準技術研究所

© Hitachi, Ltd. 2021. All rights reserved.

地球温暖化・都市化にともない、大規模災害の被害が増大

被害額1900億円*1



熊本県(2018)
地震による土砂災害・水災

被害額2.2兆円*2



東日本台風(2019)
大雨・河川氾濫

被害額2.8兆円*3



米カリフォルニア(2019)
山火事

* 写真はイメージであり、上記災害によるものではありません

*1 熊本県ホームページ「平成28年熊本地震 土砂災害の被災状況について」
*2 国土交通庁(2020年発表の暫定値)
*3 Insurance Journal試算

空撮映像を用いた災害状況の把握による災害対応の支援

災害対応の流れ

平時の構え



応急対策
(初動期)



応急対策
(初動期以降)



復旧・復興

本技術がめざす価値「被害状況の把握」

【初動期】すぐにたどり着けない現場状況を、迅速・安全・詳細に把握

【初動以降】道路などが寸断されてる中での被災状況の把握

【復旧】広範囲にわたる被害状況を効率的・網羅的に把握

■ 空撮映像の活用が期待

ドローンによる
撮影映像A



低空ヘリによる
撮影映像B



技術課題

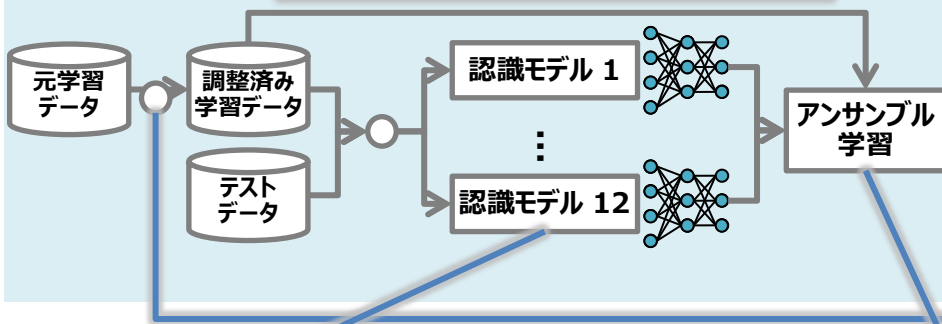
- ・認識対象がさまざま
(「地滑り」などの災害属性、
「車」などの物体)
- ・AIの学習に必要なデータが
少ない

*図出典：TRECVID2020 Test Dataset⁵⁾

空撮画像内に大小さまざまに存在する複数の災害状況・モノを認識

【特徴1】複数の災害状況・モノを認識可能

【導入手法1】複数の認識モデル



水たまり:0.90
川:0.83
家屋倒壊:0.78
車:0.03
...

認識対象:5カテゴリ、32クラス

損害	各種被害、洪水、地滑り、橋梁崩壊、家屋倒壊、煙/火事
環境	泥/汚れ、草、溶岩、岩、砂、低木、雪/氷、木(林)
インフラ	橋、ビル/建物、ダム/堤防、ガス管/水道管、電線、線路、電波塔、給水塔、道路
乗り物	飛行機、船、自動車、トラック
水	洪水、湖、海、水たまり、川

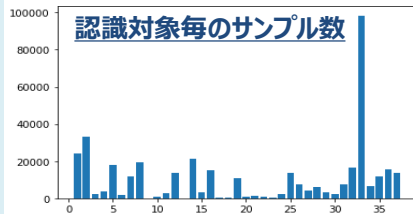
【特徴2】小さなモノの認識

広範囲の映像から、車・船など小さい物体の認識



【特徴3】学習サンプルが少ない災害状況の認識

サンプル数に偏らない認識モデルの学習



【特徴4】誤った情報を含む学習サンプルに対応した学習

人間でも判断困難な災害状況を学習



3人でつけた正解情報
A{川、泥、橋、ダム}
B{川、橋梁倒壊}
C{川、橋梁倒壊、岩}

国際ワークショップTRECVID 2020でトップレベル精度を達成

- TRECVID: アメリカ国立標準技術研究所が主催するワークショップ (TREC Video Retrieval Evaluation)
2001年の黎明期からある高い権威を持つ競争型ワークショップの一つ
- DSDI : Disaster Scene Description and Indexing、災害シーンの記述とインデキシング
- 学習・評価用データセットの詳細
 - 正解情報あり画像 : 約4万枚、複数の人が正解付け
 - 正解情報なし画像 : 約60万枚
- テスト映像
 - 1825個のクリップ映像 (1クリップ約10秒程度)



実用に向けた進化 1 : 極小物体検出

- ・被害状況を迅速に把握するため、空撮映像から検出対象の位置と数を自動算出 (※下図は車を検出)



実用に向けた進化 1 : 極小物体検出

- ・被害状況を迅速に把握するため、空撮映像から検出対象の位置と数を自動算出 (※下図は車を検出)



*出典：国土地理院撮影の映像より図を使用し、検出結果を重畳したもの

実用に向けた進化 1 : 極小物体検出

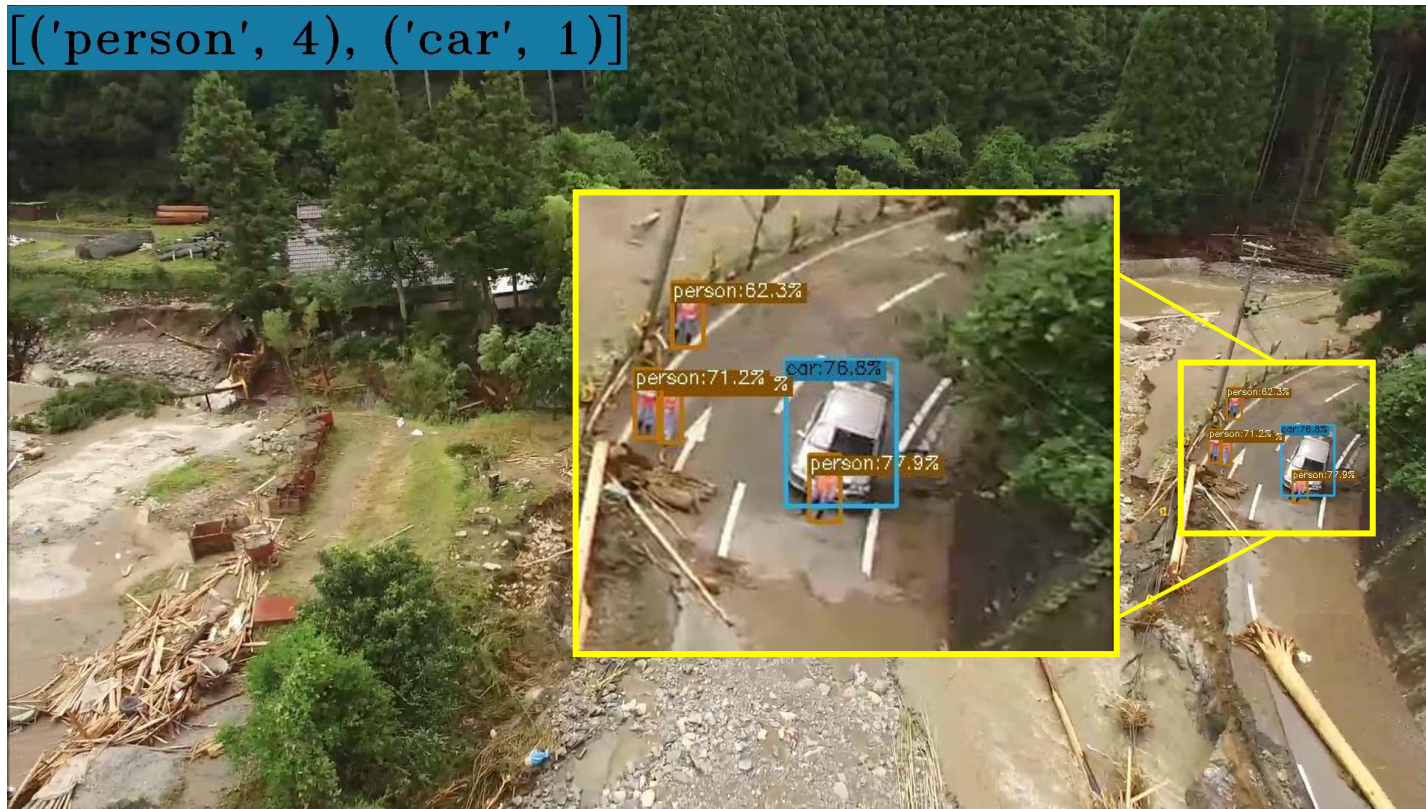
- ・被害状況を迅速に把握するため、空撮映像から検出対象の位置と数を自動算出 (※下図は人・車を検出)



実用に向けた進化 1 : 極小物体検出

- 被害状況を迅速に把握するため、空撮映像から検出対象の位置と数を自動算出 (※下図は人・車を検出)

[('person', 4), ('car', 1)]



*出典：国土地理院撮影の映像より図を使用し、検出結果を重畳したもの

実用に向けた進化2：災害位置の同定

- ・土砂崩れ(landslide), 家屋倒壊(rubble), 道路や橋の流失(washout)など、空撮映像における災害状況を自動検出し、場所も自動的に特定

[('rubble', 3), ('washout', 2), ('landslide', 1)]



[('landslide', 1), ('washout', 1)]

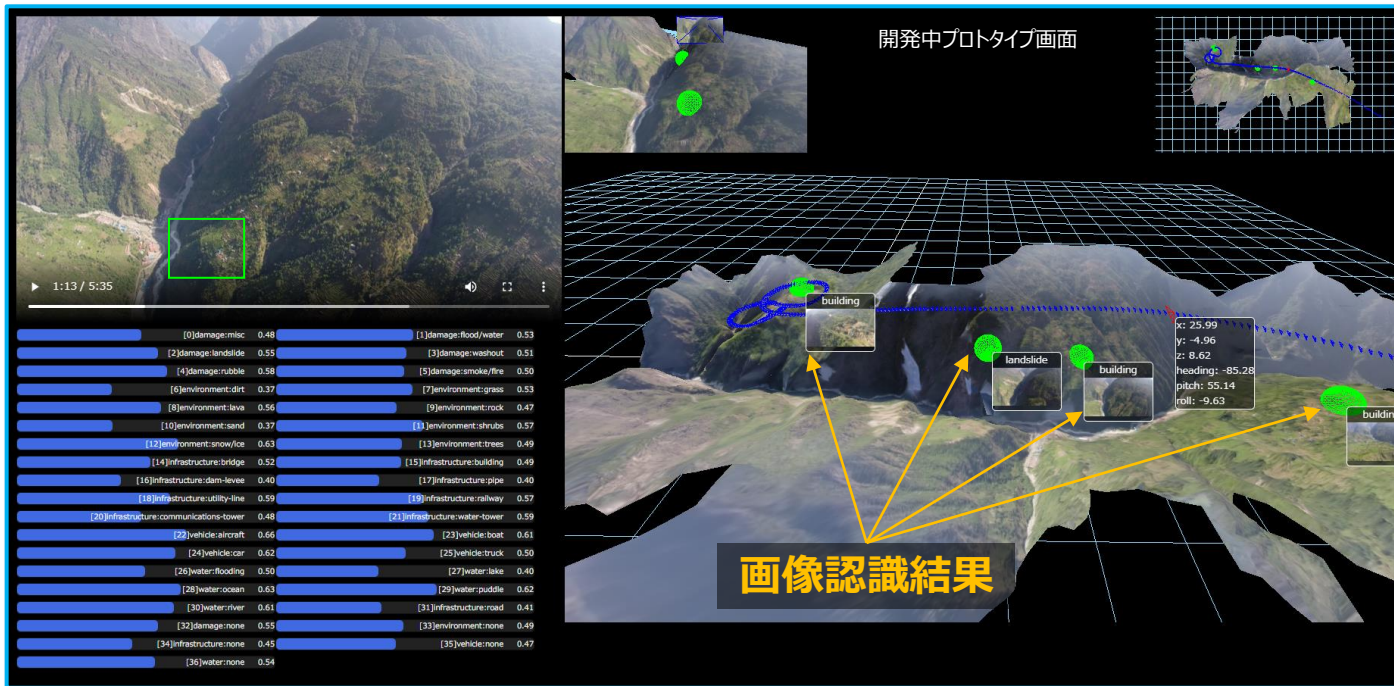


*右図出典：国土地理院撮影の映像より図を使用し、位置同定結果を重ねたもの
*左図出典：TRECVID2020 Test Datasetより

3D地図上に画像認識結果をマッピングすることで 広域の被災状況を短時間で直感的に把握可能

映像解析AI

3D地図生成



日立の技術連携で国土強じん化のための課題に対応できると考えております。

- ・被災建屋や避難所他に設置した監視カメラの画像を活用し、行方不明者の捜索が可能



※商品化済

- ・物資輸送の効率化
- ・平時の交通障害検知



- ・ドローン、衛星、カメラを活用した港湾における被災状況把握
- ※衛星画像提供ソリューションと連携



DigitalGlobe社衛星

※商品化済

映像解析システム



- ・ドローンやヘリコプターから撮影された空撮映像に対し、**災害状況を高精度に解析できるAI技術を開発**
- ・アメリカ国立標準技術研究所(NIST*¹) が主催する、歴史ある映像解析の国際ワークショップ「TRECVID 2020」の災害映像解析タスク**DSDI*²**にて**世界トップレベルの認識精度を達成**
- ・本技術を活用した災害対策の社会イノベーションを推進し、自治体や設備保守、保険会社などのパートナーさまとの協創を通じて、**レジリエントな社会、人々の安心・安全な暮らしの実現に貢献します。**

*1 NIST: National Institute of Standard and Technology

*2 DSDI: Disaster Scene Description and Indexing

Hitachi Social Innovation is

POWERING GOOD

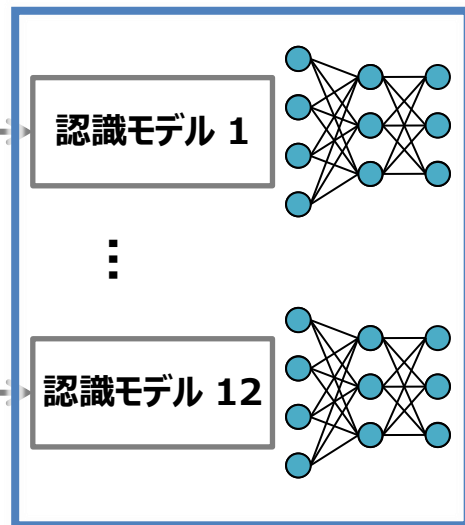
世界を輝かせよう。

HITACHI
Inspire the Next 

特性の異なる複数のネットワークを利用。注目すべき領域を判断する機構を用いることで画面全体のシーンも小物体の認識も可能

【導入手法 2】異なる領域を注目する複数の認識モデル

入力画像



モデル 1 の出力

家屋倒壊 0.5
草地 0.8
林 0.9

⋮

モデル12の出力

車 0.95
...



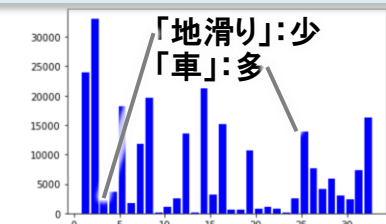
最終出力

家屋倒壊 0.7
草地 0.66
林 0.81
車 0.85

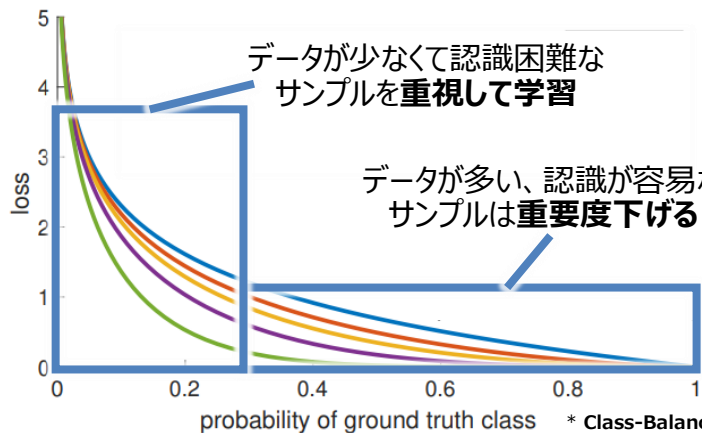
学習サンプル数に応じて、重視する認識対象をバランスさせることで、レアな認識対象も高精度に認識可能

【導入手法 3】

データ数に応じた重み付け手法



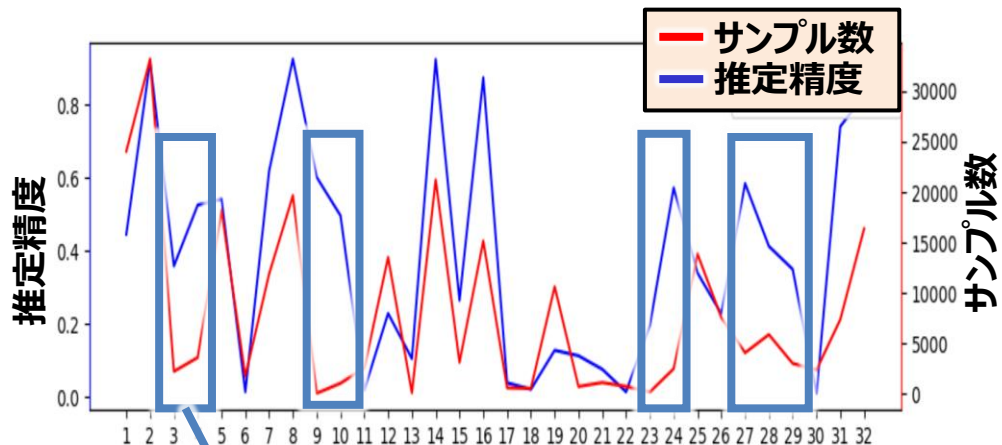
【課題】
データが少ないモノ・災害
状況は軽視されて
認識精度が低くなる



* Class-Balanced Focal Loss [CVPR 2019]
* Reduced Focal Loss [Arxiv 2019/03]

■ 効果

サンプル数が少ないクラスも高精度に認識可能に



* 図出典：TRECVID2020 Test Datasetより

誤りを含むデータの影響を軽減した学習によって、
人でさえ認識が困難な属性も認識可能に

学習画像



*図出典：TRECVID2020 Test Dataset

正解情報

Aさん{川、泥、橋、ダム}

Bさん{川、橋梁倒壊}

Cさん{川、橋梁倒壊、岩}

正解付けた人によって
判断にばらつき

学習

推論結果と正解のばらつきによって、
学習で混乱が発生してしまい、
精度の高いモデルができない



橋がある？ ない？

【導入手法 4】

正解ラベルの数による変換

変換済み正解情報

[1.0, 0, ..., 0.66, 0.33, ...,]

川

橋梁倒壊

橋

学習

正解ラベルのばらつきを
その画像が有する曖昧性
としてAIが学習

