Improving the Speed and Efficiency of AI-Enabled Damage Assessment in Insurance

Authors:

Anant Nawalgaria, Staff ML Specialist Engineer, Google
Johannes Kuhn, Sr. Technical Data Specialist, Munich Re
Dr. Ulrich Nöbauer, Sr. Technical Data Specialist, Munich Re
John McSpedon, ML Engineer, Taktile

Special thanks:

Raviteja Vemulapalli, Research Scientist, Google
Genquan Duan, Sr. Software Engineer, Google
Dr. Henna Karna, Global Insurance, Google

Executive Summary

The combination of high-resolution aerial imaging and artificial intelligence has practical applications for multiple use cases, such as manufacturing defects detection, environmental impact studies and planning, urban vegetation mapping, and precision agriculture. Now, as major natural disasters have increased, the insurance industry is using aerial imagery and AI for damage assessment. This is part of an expanded focus on post-event analysis that uses data, AI, and analytics to understand risk mitigation. Aerial imagery and AI can help address varying spectral bands of resolutions and other factors that the human eye cannot easily connect.

For example, Munich Re has developed a suite of tools and products to help insurers and their customers after a disaster like a hurricane or tornadoes. These products use AI to detect and assess roof damage from high-resolution aerial imagery captured shortly after the event. Because of the potential high impact of these models, Munich Re is constantly working on optimizing their predictions. Also, while current visual model training processes deliver high accuracy, they also require a great deal of human labeling, and iterations can be costly and lengthy.

A method of machine learning (ML) called semi-supervised learning (SSL - specifically Semi-Supervised Semantic Segmentation with Cross-Consistency Training) has demonstrated it can reduce the cost of vision model training by 50%. By using Google Cloud TPUs(TPU V3-32) for semi-supervised learning training it is possible to reduce the training time by 92% and reduce the cost of model training by 33% for Munich Re's use case. SSL learns the features on unlabeled data and then learns the real model on features extracted from the labeled data. Selecting this ML method also has sustainability implications, as reductions in compute consumption can be directly translated into carbon footprint reductions for the company's Environmental Sustainability Goals.

This paper explains how Munich Re, Google, and Taktile engineers explored and evaluated different SSL approaches and investigated using them on Google Cloud Tensor Processing Units (TPUs). It also provides evidence of how Google Cloud Visual Inspection AI allows almost anyone, even resources with very rudimentary ML expertise to get accurate results in hours, instead of months. Its trade-off is that at this point of time, the model is a blackbox, which can be used but not manually fine-tuned/customized outside the service. Finally, it gives an example of how the Taktile Platform can be used to embed the outputs of ML models, such as Munich Re's damage assessment models, in critical business decision processes. An example would be whether a claim can instantly be settled or whether a claims adjuster must be sent to the scene.

It's important to note that companies like Munich Re (due to sensitive data) are often bound by local and regional legal frameworks that govern cloud usage. These constraints can prevent the full potential of cloud technology being used in digital transformation or machine learning modeling projects. Therefore, Google Cloud is launching a Sovereign Cloud solution, which is designed to meet the requirements for security, privacy, and digital sovereignty in Europe, without compromising on functionality or innovation. Such a solution will be very useful for use cases like those described in this paper.

Introduction

“Our mission at Google is to help the insurance industry provide the gold standards for data, analytics, AI, and ML. Working with reinsurer Munich Re to leverage AI in innovative ways not only aligns with that mission, but it highlights the value and breadth of our AI/ML capabilities. With Google technology, working together with Munich Re we are able to discover and use insights across multiple use cases,” said Dr. Henna Karna, managing director of global insurance and risk management at Google Cloud.

When a major natural disaster strikes an area, a primary carrier faces the challenges to quickly and efficiently assess a large amount of damaged properties. These processes are often slow because resources like loss adjusters are scarce, access to the affected areas is difficult, or both. Munich Re has a solution called CatAI that addresses resource and access issues by combining high-resolution imagery and machine learning methods for automatic assessment of the severity of property damages.

In cases where a more detailed damage analysis by a loss adjuster is necessary, Munich Re’s Remote Inspection Solution provides a quick and comfortable solution.[1] An application on a smartphone views live video feed, records damage, obtains the client location via geo coordinates, and writes a report–all from one dashboard. This Remote Inspection Solution allows insurers to reserve adjusters for complex losses where their skills are truly needed.

The building blocks of Munich Re’s solutions are expertise in property damage assessment and its application to insurance, efficient processing of aerial imagery, and the creation of computer vision models. All three elements have to be interlinked to achieve superior results. The respective computer vision models are trained using supervised semantic segmentation, whereby a dataset consists of images and their corresponding pixel-level class-specific annotations. For each pixel in an aerial image, the model assigns a class, depending on the use case.

The current set of labeled images allows training of a state-of-the-art model, but the creation of these images is very cost and resource-intensive. Benchmark academic datasets in semantic segmentation typically contain thousands or tens of thousands of images. Major labeling companies charge upwards of $3 USD per image for labeling these types of photos. This also implies that the process of identifying the right classes and eventually changing the classes later on are lengthy and costly iterative cycles. Examples of when labels need changing include a reevaluation of the ground truth because of new laws or the expansion of the product (for example, adding a new class for fallen trees next to the building in order to match new insurance product coverages).

Over time, the question has become “Is there a faster and more cost-effective method of training computer vision models?” In this paper, we share our investigation of Google Visual Inspection AI, semi-supervised learning (SSL) methods for classification and segmentation, and TPUs to determine if they could shorten training time and lower costs.

[1] Next-generation claim reporting is here – introducing the Remote Inspection video tool

Using Google Cloud Visual Inspection AI for fast and accurate semantic segmentation

As one of our initial experiments, we looked into whether Google Cloud Visual Inspection AI[2] could improve the labeling performance (that is, mIOU) of the existing models, using the CatAI segmentation model as a reference. This model was built over the course of several months and involved significant data preprocessing and data engineering effort. Visual Inspection AI is a managed service that is designed with many components to be label efficient. It provides segmentation-based cosmetic solutions to detect defects in images and to predict pixel-level segmentation and image-level defect classification results with high accuracy, with the tradeoff of the model being a blackbox, which can be used but not manually fine-tuned/customized outside the service (at this point of time).

Feature highlights of segmentation-based cosmetic solutions include:

- Model configurations for fine-tuning model performances easily
- Inline active learning to iteratively annotate images and reduce labeling cost
- Docker containers for scaling deployment on-premises easily
- Easy-to-use UI features with labeling, training, evaluation, and predictions

Compared to general segmentation models, the main advantages of the Visual Inspector AI training algorithms are:

- State-of-the-art segmentation backbones with different learning capacities
- Automatic data synthesis and oversampling to handle the issue of few examples
- Automatic mining of hard examples to handle the high imbalance of defect and non-defect images
- Support image tiling to handle high resolution images
- Post-process segmentation pixel prediction results to image classification prediction results for easily determining whether images are defective or not

Once the Visual Inspection AI service is used, the output stays with the subscriber. In other words, customers can access their own models, and no other customer has access. In addition, Visual Inspection AI does not use one customer's models to train another customer's models. Trained model weights are included in the whole solution artifacts. Currently, the solution artifacts can be downloaded as visual inspection containers. The containers can make sure predictions are correct.

[2] Visual Inspection AI

"The hypothesis was that Visual Inspection AI had the potential to significantly reduce the time spent creating, tweaking and training custom models with comparable or even better performance," said Dr. Maximilian Eber, chief technology officer at Taktile. In industry use cases (for example, manufacturing), Visual Inspection AI has been able to train and deploy AI models to detect production defects automatically and rapidly.

We trained the Visual Inspection AI models by keeping the test set constant and using various proportions of the training data. For all these experiments, we used the "higher capacity" configuration of the default Visual Inspection AI model. Here are our various experiments and their corresponding results.

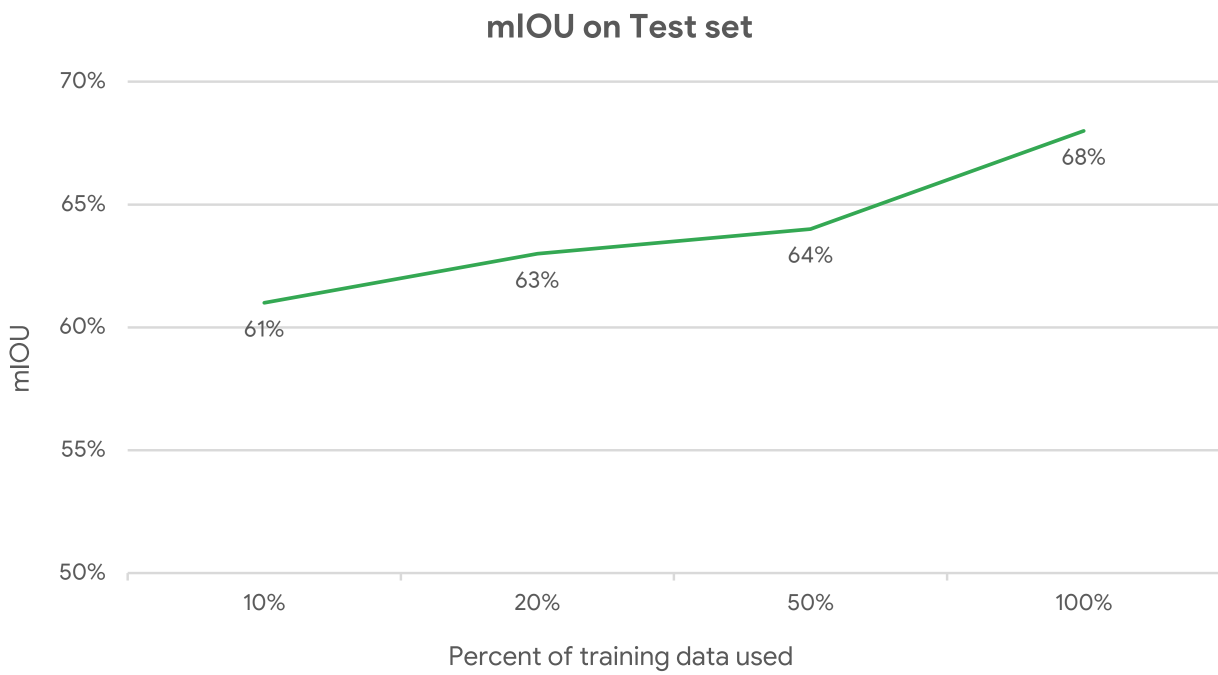


Figure 1: Visual Inspection AI test result.

The results of the Visual Inspection AI experiment (see Figure 1) were comparable to the CatAI baseline. By using a limited amount of labeled training examples very efficiently, Visual Inspection AI therefore provides a good performance point for getting an idea of how difficult the target domain is. The main advantage of Visual Inspection AI is that it delivers results quickly and efficiently, and it is not necessary to create, train, and deploy models. In our case, however, Munich Re had invested several months of domain expertise into creating the CatAI segmentation model.

Additionally, the whole experiment with VisualAI and the segmentation data set collected by Munich Re concluded in a few weeks. We estimate that if we had used an automatic Active Learning Segmentation tool in Visual Inspection AI, the labeling process would have finished in a few hours. The process of launching the service and visualizing the results was fairly easy with only a few lines of code.

Using Semi-Supervised Learning for Training Computer Vision Models

The primary objective of our collaboration was to evaluate if CatAI model-training resources (Figure 2) could be shifted from human labeling to computational work for a net savings in time, cost, or labeling flexibility. Specifically, we investigated recent advances in semi-supervised learning for image classification and segmentation as they applied to computer vision and our aerial imaging dataset. A modern paradigm in AI research is to train a large, general-purpose "upstream" model a handful of times at most, and fine-tune it for many smaller, label-scarce downstream tasks.

Figure 2: Output of damage segmentation model developed and used by Munich Re.

Likely the best known instance of an upstream model is GPT-3. One 175-billion parameter text generation model was trained across a very large body of text. After this initial expensive training, countless natural language processing sub-tasks and business models can be solved using the model as a backbone. We attempted a similar solution, but for computer vision models, with a much more modest parameter count of ~50 million.

The main attribute of "semi-supervised learning" is in its name. It is not entirely supervised; in other words, not all of the data it uses for model training requires human hand-labeling. In the case of text data, the semi-supervised tasks are intuitive. The model is fed a sentence with some words removed and trained to predict the missing words, such as "A person went to the ____ to buy a ___ of milk." An effective semi-supervised task for image data is much less obvious.

Researchers have converged around a family of related methods only just recently. The exact technical details of different approaches vary, but they generally share the idea of taking an unlabeled image from the target domain (in our case, an aerial image of buildings) and transforming it into two separate augmented views (Figure 3). These transformations are the composition of randomly sampled crops, reflections, colorizations, saturations, and so on. The model is then trained to embed all of these augmented views so that views from the same source image are near each other, and views from all other images are far away.

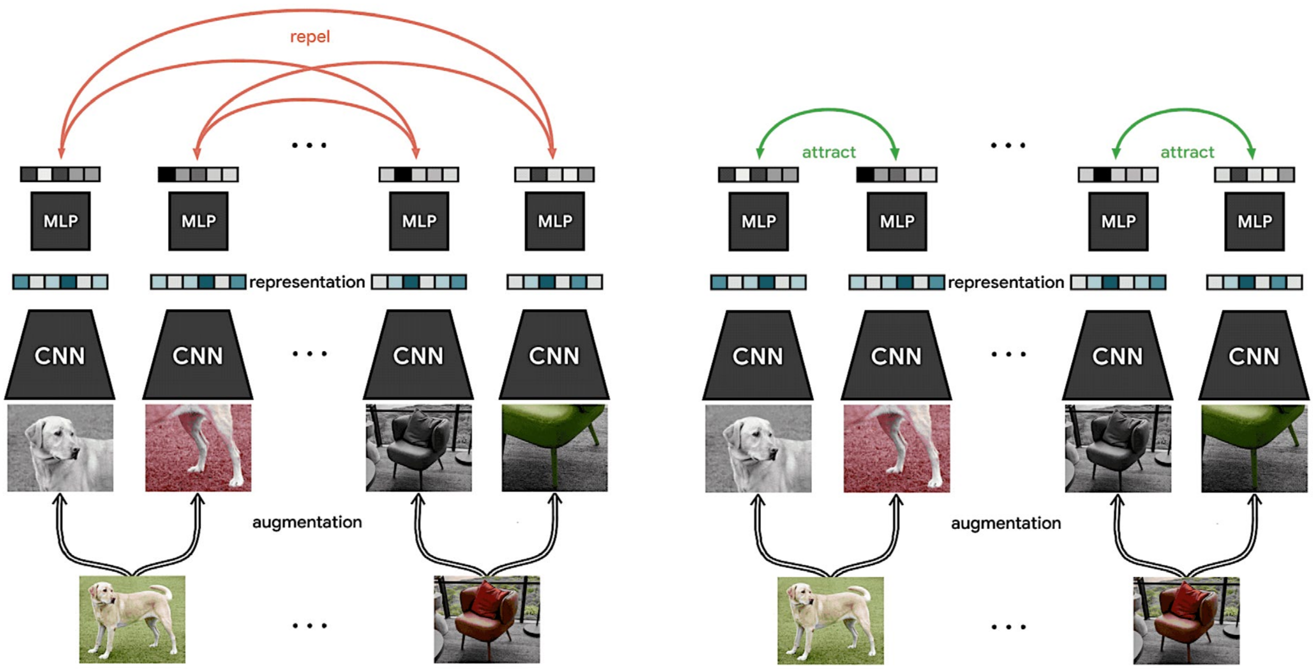


Figure 3. Transformation of an unlabeled image into two separated for training.

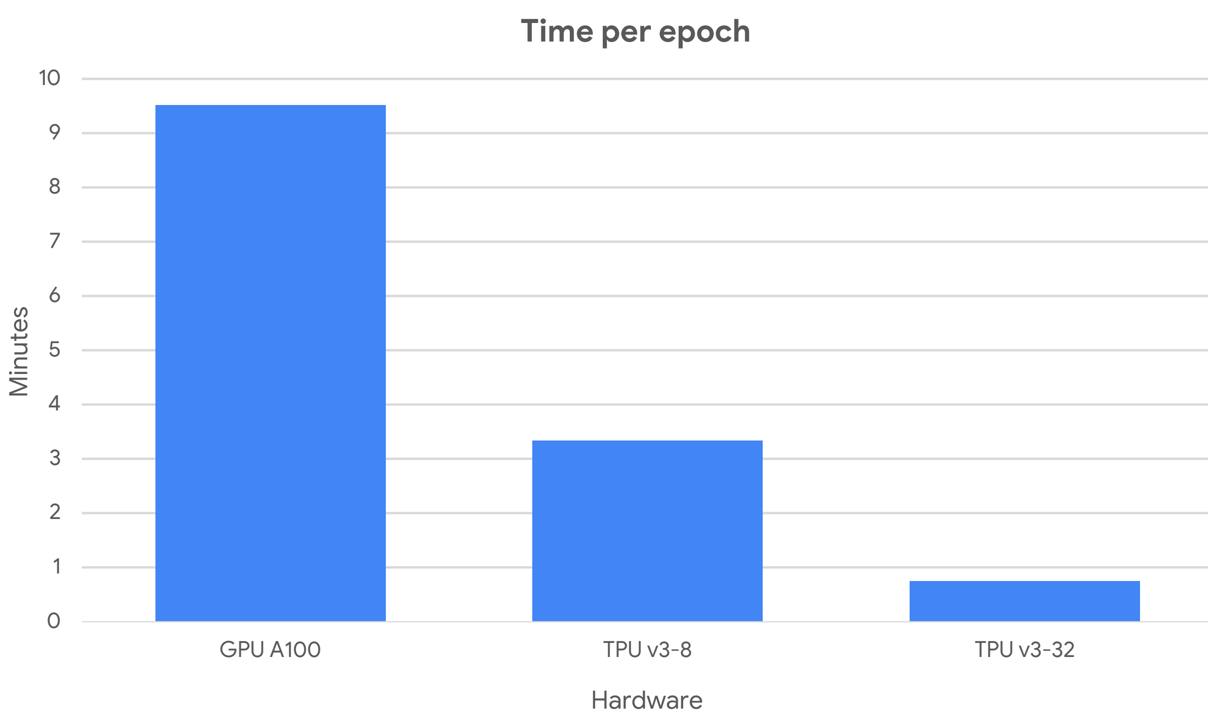
For our semi-supervised learning experiments, we explored SimCLR for classification and the Cross Consistency Training (CCT) for segmentation.

Using SimCLR for Semi-Supervised Learning for Classification

For our experiments with image classification, we used the SimCLRv2 codebase.[3] Other recent semi-supervised learning methods for representations are available through VISSL or paper-specific repositories.[4] We chose the SimCLR codebase because it matched our preferred deep learning framework of TensorFlow.

“For hardware, we selected tensor processing units (TPUs), which are application-specific integrated circuits (ASICs) developed by Google Cloud to accelerate machine learning workloads. Our CatAI baseline model was written in TensorFlow (with existing data pipelines and evaluation code optimized for TensorFlow), and we were excited to try out Google Cloud TPUs,” explained Thilo Horner, senior project manager of innovation at Munich Re.

TPUs have large on-chip memory and high bandwidth interconnects that make batch sizes in the thousands very achievable, even for higher parameter count models. We tested SimCLR against different hardware available on Google Cloud Platform. We found that running TPU v3-32 was 33% less expensive (in US dollars) per epoch than our graphical processing unit (GPU) A-100 baseline (see Figure 4), which are also popular for deep learning.



[3] https://github.com/google-research/simclr/tree/master/tf2

[4] https://github.com/facebookresearch/vissl

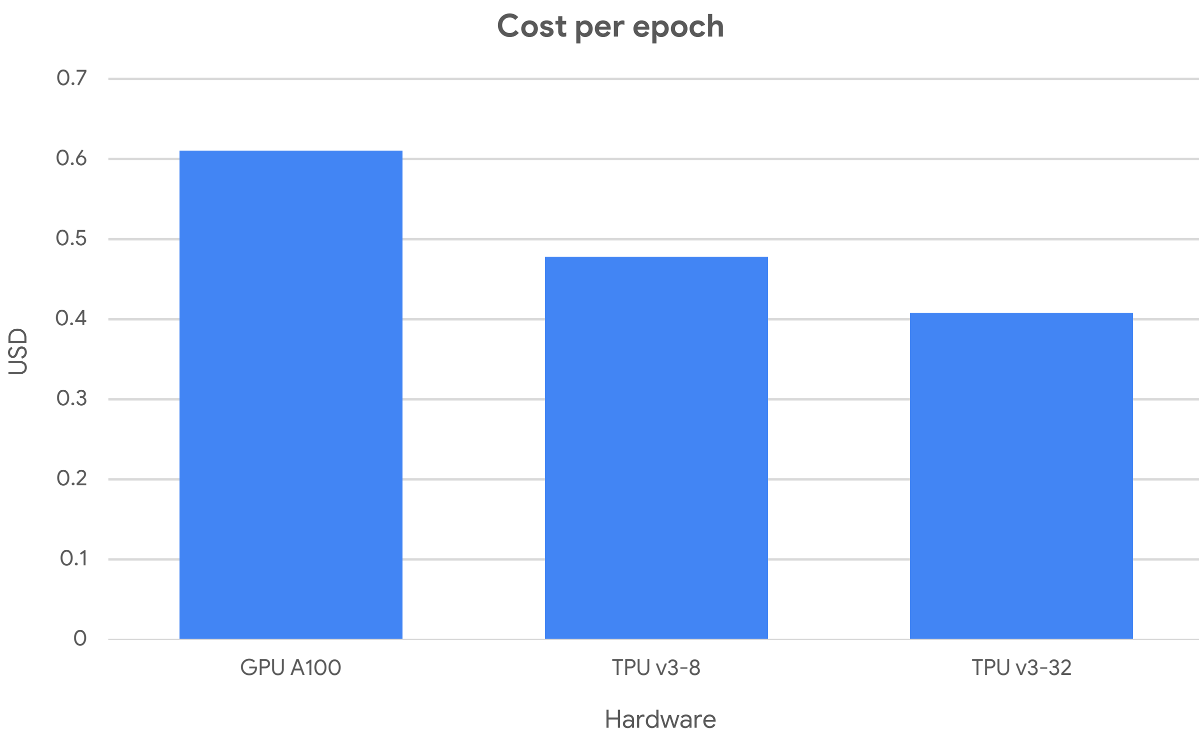


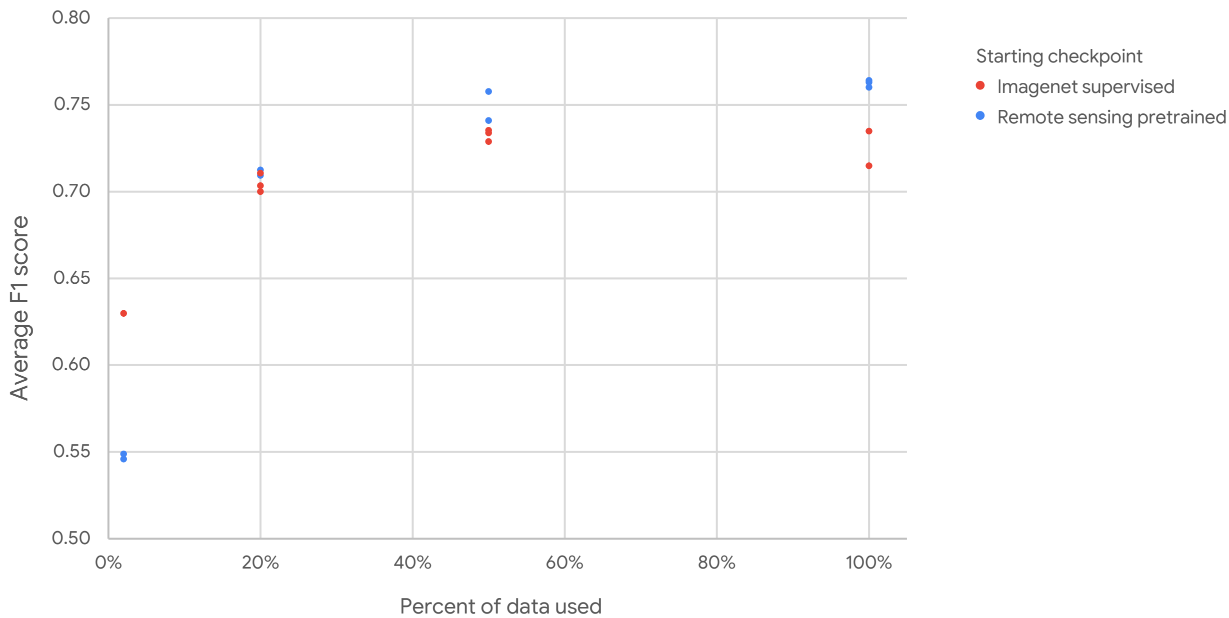
Figure 4: GPUs versus TPUs in terms of time and cost.

More significantly, training ran 12.5 times as fast in walltime. This is arguably a greater cost savings in terms of researcher or employee time. Model training cycles that originally ran overnight completed in hours, allowing our engineers to iterate more quickly in fewer work days overall.

We ran a subset of experiments based on the original SimCLRv2 paper to assess whether this SSL approach improves label efficiency when applied to the CatAI use case. Namely, we chose a resnet-50 architecture and compared the performance of two starting checkpoints. One is “imagenet supervised”; we used traditional transfer learning by pretraining on imagenet. The other is “remote_sensing_pretrained”; we performed SimCLR unsupervised training on random aerial images, and used this backbone as our starting checkpoint.

We compared the performance of these models when they’re trained for convergence on 2%, 20%, 50%, and 100% of the available training and validation data (in all cases we test on 100% of the test data). As expected, pretraining on aerial imagery improves model performance in the “label scarce” domain of 2% data (Figure 5), We also eked out higher performance in the full data domain of 100%. Results are mixed or slightly negative at 20% and 50% of available labels.

We don’t have a compelling hypothesis for these results. For future experiments, we plan on increasing the input size of input images from Imagenet standard 224 x 224 pixels to the 384 x 384 pixels size used by CatAI and filtering our pretraining aerial images so that they always contain buildings.



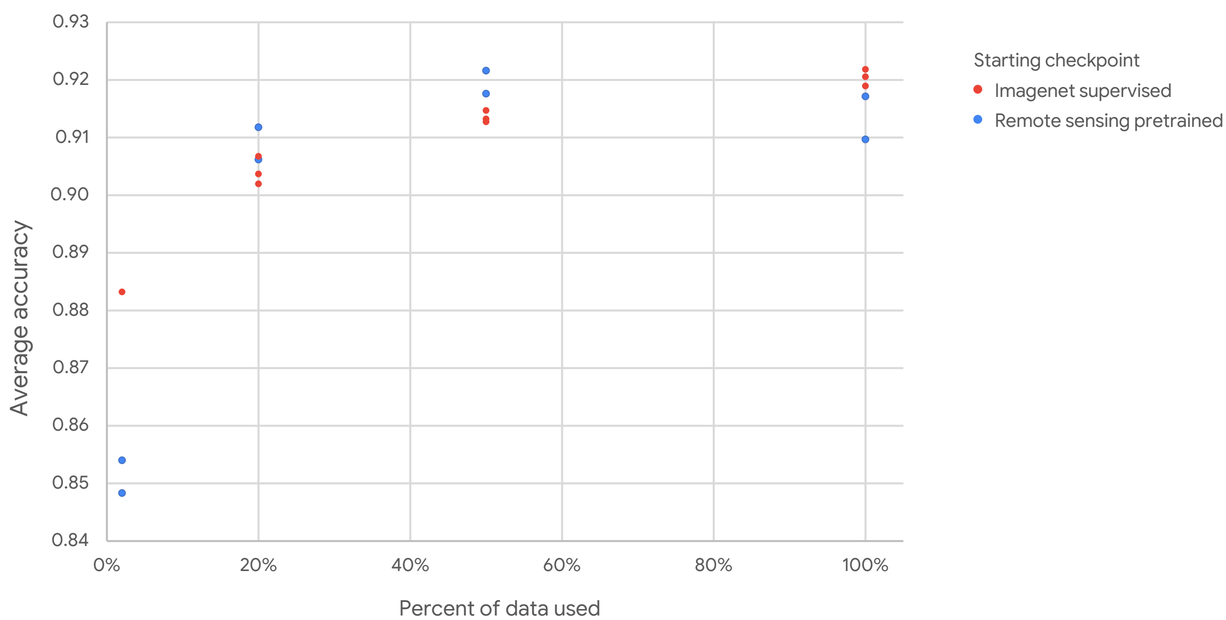


Figure 5: The performance of supervised and random unsupervised models when they're trained for convergence on 2%, 20%, 50% and 100% of the available training and validation data.

Using CCT for Semi-Supervised Learning for Segmentation

We evaluated different state-of-the-art self-supervised learning approaches for semantic segmentation before ultimately choosing the Cross-Consistency Training (CCT) approach proposed by Yassine Ouali, Céline Hudelot, and Myriam Tami in "Semi-Supervised Semantic Segmentation with Cross-Consistency Training." [5]

CCT uses supervised learning to train a shared encoder and a main decoder with the available labeled examples. To use unlabeled examples, a consistency between the main decoder predictions and those of the auxiliary decoders is enforced. The inputs are different perturbed versions of the encoder's output, and consequently, improve the encoder's representations.

The goal of our experiments was to evaluate the performance of the CCT approach within our domain and to evaluate label efficiency. We did not strive to develop the best possible model with extensive tuning of hyperparameters.

Figure 6 shows the results of the experiments. The green line represents the baseline trained using only the supervised part of the CCT approach with 100% of the labeled data. We used this as a baseline in our experiments to keep as many parameters as possible stable during our experiments.

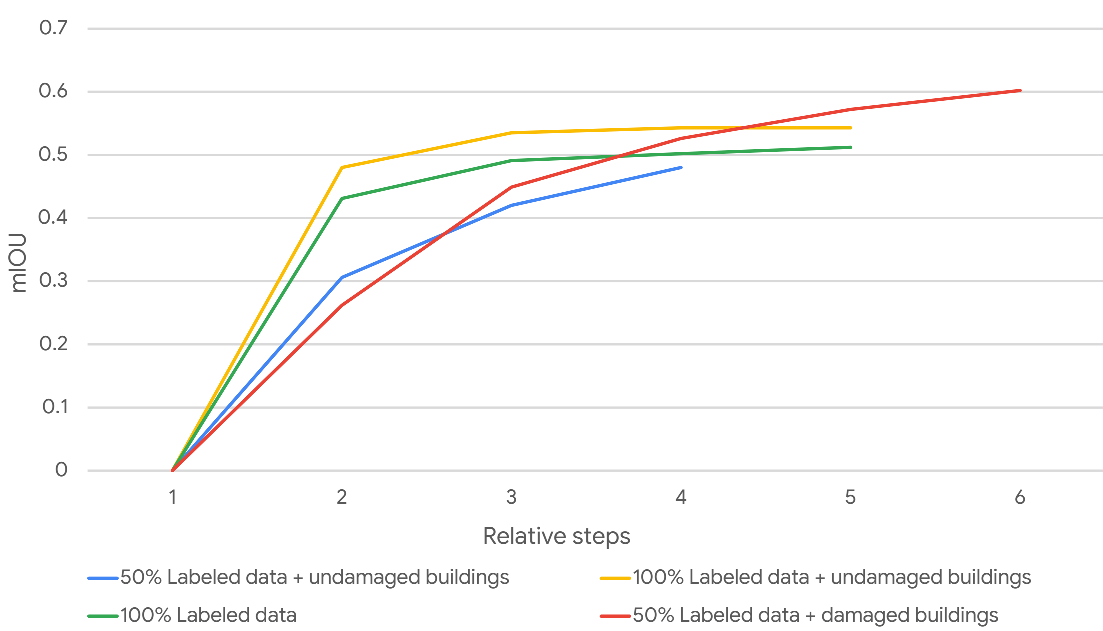


Figure 6: Comparison of CCT results for labeled, unlabeled, and half-labeled images.

[5] Yassine Ouali, Céline Hudelot, Myriam Tami, "Semi-Supervised Semantic Segmentation with Cross-Consistency Training," 2020. Cornell University.

When using 100% of our labeled data and using the same amount of images of undamaged buildings, the overall performance of the model can be improved (yellow line). The results show that the auxiliary unlabeled images contain additional useful information. The costs of additional unlabeled images is minimal. Further experiments in the future must show to what degree adding more images is helpful.

For the labeling efficiency tests, 50% of the labeled data was used and supplemented with either the same number of undamaged (blue line) or damaged buildings (red line). When using only undamaged buildings as auxiliary data, the model achieved approximately 90% of the performance of the baseline. If damaged buildings are used the performance increases by 15%. This highlights two important facts. First, the domains of damaged and undamaged houses are clearly separated from each other. The use of damaged buildings has a significant impact on the result. Second, the CCT approach can significantly improve label efficiency. The concrete results depend on the proximity of the domain used for the unsupervised part. From the economic side, it must be taken into account here that obtaining images from different domains may also have different costs.

Art of the Possible: Turning Model Predictions into Decisions on Taktile

After the new computer vision models are trained, it is time to use their predictions (outputs) as inputs in a real-life decision process. In our example, the Munich Re CatAI damage classification data can be used to support decisions that an insurer has to make to effectively triage a reported property damage. The scope of decision ranges from opening or disregarding the claim to settling the claim instantly, using Munich Re's Remote Inspection Solution to remotely collect more evidence or sending a claims handler to also enable mitigation measures to prevent higher losses. In addition to model outputs indicating the severity of the damage, this decision has a lot of other input factors that are related to:

- The underlying insurance contract, such as whether a claim is actually covered
- Business considerations, such as whether sending a claims handler to address damage volumes under a certain amount is too costly
- The strategic goals of an insurer, such as increasing customer satisfaction in a specific customer segment.

These factors are typically expressed as heuristics or business rules in a tree-based decision structure, which also includes the outputs of the computer vision models and the digital and traditional process capabilities of the insurer. In the insurance industry, this is typically called a "claims decision flow." Insurers must constantly adjust this flow to improve their bottom lines, meet new strategic goals, or react to new business realities where the standard claims decision process is not fully applicable, such as a natural catastrophe.

To build, manage, and optimize these decision flows, insurance companies like to rely on decision engines like the Taktile Platform, which runs natively on Google Cloud. This platform helps insurers manage their complex decision flows by providing an easy-to-use workflow tool.

Subject matter experts, such as claims professionals, can combine the outputs of machine learning models deployed on platforms like Google Cloud Vertex AI, augment it with the Taktile Platform, edit business rules in a no-code environment, and deploy it without the help of engineering. Firms are then able to update their decision flows at a much lower cost and significantly reduce time-to-market, which can make a big difference for customers who require fast claims settlement.

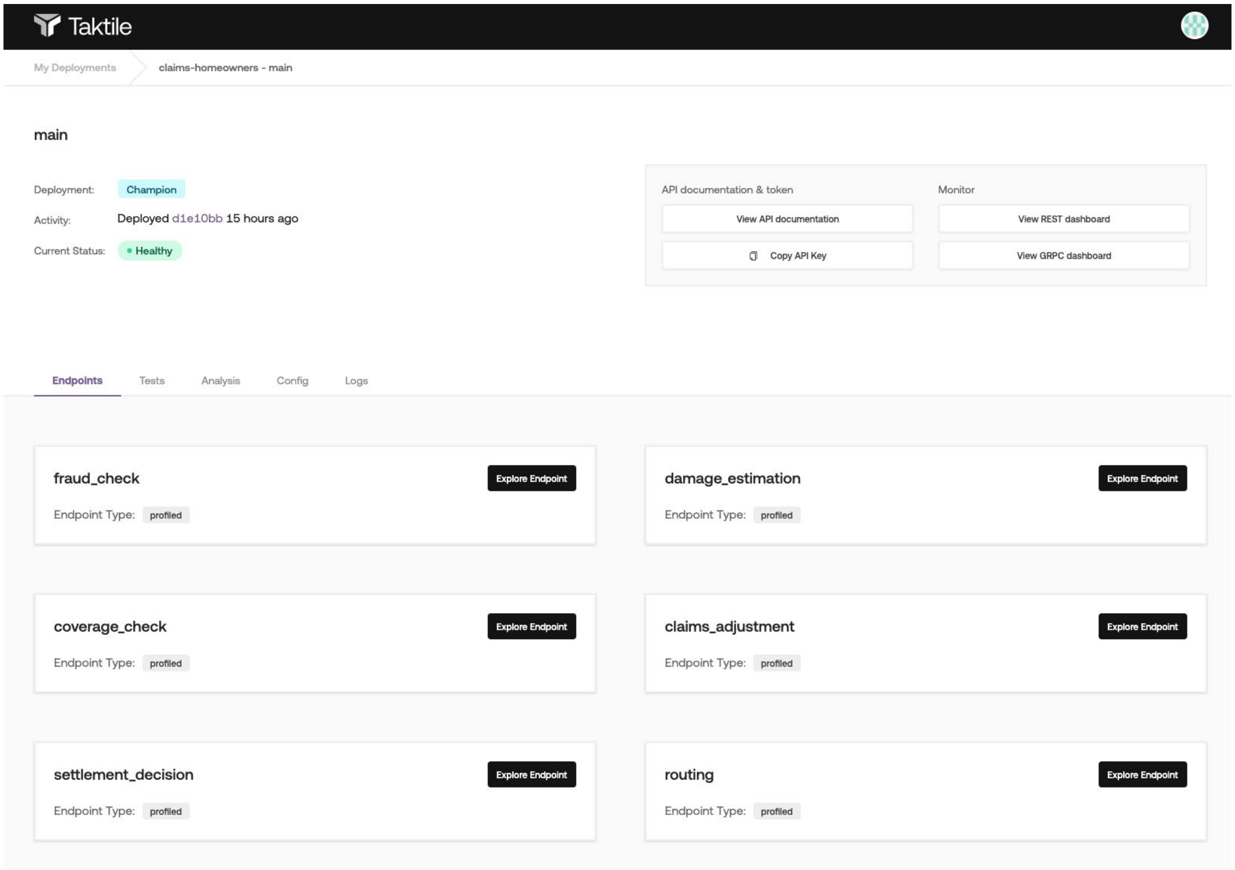


Figure 7: The Taktile Platform user interface.

Conclusion

We experimented with different options for determining if there were faster and more cost-effective methods of training computer vision models that Munich Re can use to augment their current, highly accurate visual computer modeling. Our results showed great promise for several solutions and methods. All tested methods have clear advantages and disadvantages. When chosen to the best of their capabilities, each solution can shine:

Google Visual Inspection AI is an excellent option when ease of use, turnaround time, and a strongly performant model that can easily be deployed without manual heavy lifting are the most important requirements. However, you cannot tweak the approach as much as in a custom implementation and therefore might miss a few performance percentage points.

Google TPUs are the accelerators to choose if you need performance and have high demand for compute power or RAM. If your use case demands a lot of parameters or huge amounts of training data or a big batch size (SimCLR), you can really benefit from TPUs. This comes with the cost of higher implementation efforts and less publicly available resources compared to using more traditional GPUs for training.

Self-implemented semi-supervised solutions can reduce the number of necessary training examples to reach a reasonable level of performance or they can be used in addition to a more traditional supervised approach to get the absolute best performance. This freedom of choice comes with the cost of higher efforts and knowledge necessary compared to auto-ml approaches like Visual Inspection. We're optimistic these approaches will become easier in a year or two; computer vision typically sees advances made in image classification propagate to segmentation with a 6-24 month delay. Another common challenge in image segmentation is that the annotations are not very accurate along the pixels at the edges of objects, which results in naturally occurring input-dependent label noise.[6] In the future, we plan to explore techniques that make neural networks robust to such input-dependent label noise.[7]

Of course, these different approaches are not mutually exclusive and depending on the requirements and business needs, a combination might be optimal. For example, Visual Inspection AI might be a good starting point to get a good baseline and fast results. Later on, you can switch to a custom implementation on TPUs if the business case allows it.

[6] Mark Collier, Basil Mustafa, Efi Kokiopoulou, Rolphe Jenatton, Jesse Berent, "A Simple Probabilistic Method for Deep Classification under Input-Dependent Label Noise," 2020, Cornell University

[7] Mark Collier, Basil Mustafa, Efi Kokiopoulou, Rolphe Jenatton, Jesse Berent, "Correlated Input-Dependent Label Noise in Large Scale Image Classification," 2021, CVPR 2021 Open Access

About the Partners

Munich Re

Munich Re is one of the world's leading providers of reinsurance, primary insurance and insurance-related risk solutions. Munich Re is globally active and operates in all lines of the insurance business. Since it was founded in 1880, Munich Re has been known for its unrivaled risk-related expertise and its sound financial position. It offers customers financial protection when faced with exceptional levels of damage – from the 1906 San Francisco earthquake through Hurricane Ida in 2021. Munich Re possesses outstanding innovative strength, which enables it to also provide coverage for extraordinary risks such as rocket launches, renewable energies or cyber risks. The company is playing a key role in driving forward the digital transformation of the insurance industry, and in doing so has further expanded its ability to assess risks and the range of services that it offers. Its tailor-made solutions and close proximity to its customers make Munich Re one of the world's most sought-after risk partners for businesses, institutions, and private individuals.

Google

Google's mission is to organize the world's information and make it universally accessible and useful. To that end, Google Research teams aspire to make discoveries that impact everyone, and core to our approach is sharing our research and tools to fuel progress in the field. Google Cloud translates the research into value, helping customers meet their business challenges head-on.

Named a Leader in The Forrester Wave™: AI Infrastructure, Q4 2021, we enable businesses to put the best of our artificial intelligence to work confidently, knowing they have industry-leading solutions that enable them to run faster and smoother, while finding new ways to delight their customers. For the insurance industry, in particular, our mission is to provide the technology and capabilities the industry needs to help our world become more resilient to risks that are both parameterized or emerging.

Taktile

Taktile is one of the leading software providers for automated decision-making in the financial services industry. It exists to create value for organizations through smarter and safer decisions at scale and, to date, Taktile's software has been used by our insurance and banking customers to power over 20 million automated decisions in the financial services sector.