Machine Learning and Construction Safety
A CASE STUDY ON USING COMPUTER VISION FOR SAFETY SCREENING IN ENR'S ANNUAL PHOTO COMPETITION

INTRODUCTION: CRACKS NOT CATS

As consumers, we often take it for granted that typing "cat" into an Internet search engine will present the results we're looking for back to us. But if we needed to search for useful imagery for our company's safety or marketing departments, could we find it with the same level of speed and sophistication as cat pictures? What methods are available for searching through photos and videos we use for work? The solution to this problem lies in applying artificial intelligence technologies known as machine learning, often used in the consumer world, to problems in industry.

Over the last five years, amazing progress has been made in the fields of artificial intelligence and machine learning. Microsoft, Apple, Facebook, Salesforce and Alphabet (Google) have all implemented major machine learning initiatives, with excellent results. As one example, Google Translate's error rate dropped by 60% after switching to machine learning models in 2016[1].

Over the last 20 years, improvements in artificial intelligence (AI) have been demonstrated in various public competitions; IBM's DeepBlue beat Russian Grandmaster Kasparov in chess in 1997, IBM's Watson achieved victory over Jeopardy champions in 2011, and Google's DeepMind won a game of Go against a professional player in October of 2016. In the tradition of these events, this case study presents the results of an AI compared to a human to illustrate the potential value of machine learning for construction safety.

SUNDAR PICHAI, CEO of Alphabet
Google Earnings Call[2]:

"Machine learning is a core, transformative way by which we're rethinking everything we're doing."

In partnership with Engineering News-Record, we present the first-ever 2016 *Artificial Intelligence for Safety Demonstration*. We hope these results demonstrate the potential value machine learning holds for construction safety, and how these traditionally consumer-oriented technologies can now be applied to solving industrial-grade problems.

SMARTVID.IO

OUR VISION: USING AI TO HELP INDUSTRY

There is a significant amount of photo and video content created in the field; for example, it's not uncommon to generate at least 1 gigabyte of media per week on AEC (Architecture, Engineering and Construction) projects. The collection, organization, and retrieval of industrial media is a challenge. But it is also an opportunity.

ImageNet's popular image recognition competition shows computers now beating human accuracy. [4]

Ever cleverer
Error rates on ImageNet Visual Recognition Challenge, %

HUMAN LEVEL

2011 12 13 14 15
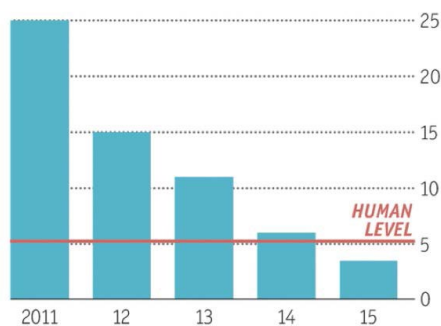
Sources: ImageNet; Stanford Vision Lab

Economist.com

Image recognition (also known as computer vision) has been one of machine learning's biggest improvements, and computers can now categorize millions of images faster and better than their human rivals. Unfortunately, even many new and improved image recognition tools are not a fit for industrial uses. Much like Apple's Siri can recognize "room service" but doesn't recognize "Room 1513-A," Google's image recognition can easily categorize "cats," but struggles to identify "cracks in concrete."

The Smartvid.io solution is a media management platform, designed with the specific needs of industrial users in mind. The application is project-centered; industrial photos and videos taken by team members using "off the shelf" equipment such as digital cameras, cell phones and GoPros are all aggregated into one place. Our

machine learning engine, nicknamed *VINNIE* (for 'Very Intelligent Neural Network for Insight and Evaluation'), then uses speech and image recognition to analyze and organize project files.

As files are uploaded, they are automatically tagged by their visual and audio content based on what the AI engine, VINNIE, sees and hears in the content. We call these tags SmartTags as no additional effort is required from the user. Since VINNIE's recognition ability is broad, users can create a list of 'tag definitions' to limit its recognition to only those things that are useful to your team. The tagging system also gives flexibility for content search and retention – think of a search able to pull thousands of variations of tags as criteria to show flexible "slices" of your media database, based on your team's current needs. Like safety...

CASE STUDY: THE 2016 AI FOR CONSTRUCTION SAFETY CHALLENGE

Engineering News Record, the #1 publication in the construction and capital projects industry, holds an annual *Year in Construction Photo Contest*[3] where readers from AEC firms submit thousands of project images to be selected for publication in the widely circulated magazine. All images are judged for artistic merit, and simultaneously reviewed by a safety expert to ensure that none

VINNIE detecting workers in jobsite photos (Photo courtesy of ENR)
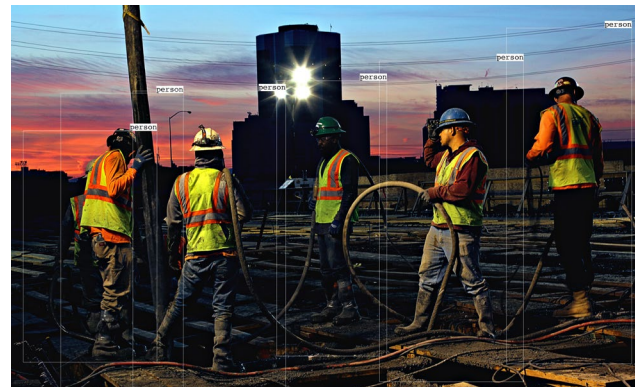
depict any hazardous situations. Detection and prevention of safety risks is one of the most common needs of industry, especially so within AEC. Approximately 1 in 5 American worker fatalities occur in construction, with jobsite deaths in 2015 rising to their highest level in seven years.[5]

For the 2016 ENR Photo Competition, Smartvid.io partnered with ENR's editorial staff to hold the first-ever *Artificial Intelligence for Construction Safety Demonstration*. In the spirit of other human vs. AI demonstrations, VINNIE analyzed all 2016 contest images alongside the traditional human safety experts. Here's how it worked:
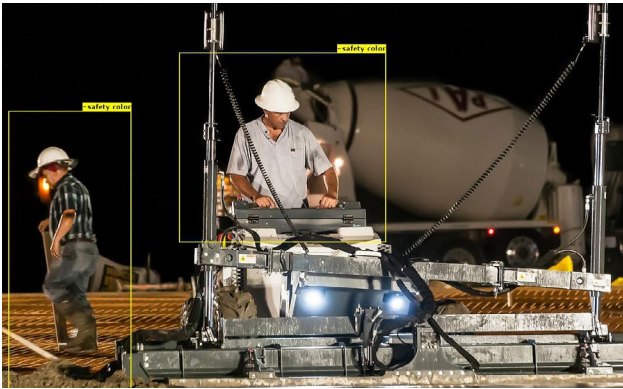
First, VINNIE automatically scanned photos to for the presence of human workers. Simply knowing that people exist in an image is a great way to flag content a human safety expert needs to review. Next, the people-containing images were automatically checked for two specific categories of safety risks: 1) workers not wearing hard hats and/or 2) working not wearing safety vests.

VINNIE detecting people missing safety vests (Photo courtesy of ENR)

The results were startling. VINNIE was able to sort through all 1,080 submissions in under 10 minutes, while the human team required over 4.5 hours. VINNIE correctly detected 446 images with people, while the human review only found 414. VINNIE also flagged 32 images containing personnel missing hard hats, and 106 images with workers missing safety-colored clothing. Human judges agreed that the implications of using a tool like VINNIE were better speed, accuracy, and

"wider coverage" of safety checks. This monitoring can provide automated flagging of potential issues, giving human experts more time for careful review.

VINNIE detecting people missing hard hats and safety vests. (Photo courtesy of ENR)
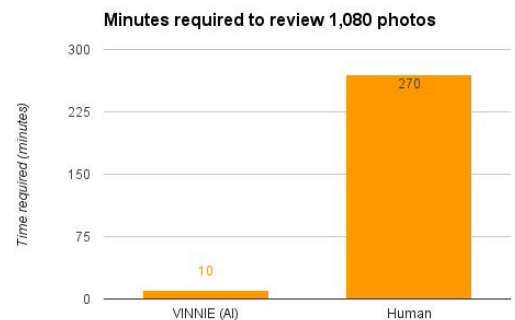

CONCLUSION & NEXT STEPS

The results of the *AI for Safety Demonstration* show the potential value machine learning has for safety, by being able to weed through thousands of photos at high speed in order to flag risk indicators for safety experts. Using computer vision to spot potential risks certainly does not remove the need for humans, however; a safety professional has years of training and can spot thousands of issues, whereas VINNIE can currently recognize only a few specific categories. However, this sort of automated field monitoring adds leverage to safety efforts as pair of "computer eyes" that never sleep, identifying risk with greater speed and accuracy than otherwise possible.

SMARTVID.IO

What comes next? First, we are working to add additional capabilities for VINNIE – beyond people, hard hats and safety vests – to new categories of potential risks.

Second, we are working hard to make VINNIE's capabilities more accessible by integrating our Smartvid.io platform into the systems you already use to store your photo and video content. Via our integrations with Procore®, Autodesk's BIM 360 Field® and other products, getting your old or ongoing photo and video content into our platform for VINNIE's review is now easier than ever, and can help your safety, marketing, and operations teams be up and running quickly.

> **MARTIN LEIK, Regional Safety Director Suffolk Construction**
>
> *"We continue to push the boundaries of what is possible on our projects, so it's exciting to test innovative technologies that could have a positive impact on the entire construction experience."*

Recognition capacity for "People", "Hard Hats" and "Safety" categories is available as part of the Smartvid.io free trial. Go to www.smartvid.io to sign up, start screening your photos, and make searching for safety risks as easy as cat pictures.

We invite your feedback at: www.smartvid.io/contact as we welcome you to join us on the journey to bring machine learning to the world of industry.

A BIG "THANK YOU" TO OUR CONTRIBUTING PARTNERS

In preparation for the AI for Safety demonstration, VINNIE expanded on previous safety training by reviewing images provided by a set of partners who are dedicated to promoting safety in construction.

ENR led the way, contributing anonymous photos from 10 prior years of photo competitions to provide material for Vinnie's training. Additional imagery was provided by the partners listed below, and others who have decided to remain anonymous. We are grateful for their help and support.

REFERENCES

1   D. Castelvecchi, "Deep learning boosts Google Translate tool," *Nature* , 27 9 2016.

2   Seeking Alpha, "Alphabet (GOOG) Q3 2015 - Earnings Call Transcript," 23 10 2015. [Online]. Available: http://seekingalpha.com/article/3596706-alphabet-goog-q3-2015-results-earnings-call-transcript.

3   Engineering News-Record, "Annual Photo Contest Archives," 2016. [Online]. Available: http://www.enr.com/photoContest/archives/.

4   The Economist, "From not working to neural networking," *The Economist*, 25 6 2016.

5   T. Ichniowski, "Construction Fatalities climb in 2015," *Engineering News-Record* , 12 2016.