

Image Recognition and the Future of Data

Anthony Wang

From predictive modeling to process automation, advancements in actuarial technology amplify the efficiency of actuaries; consequently, the productivity of a single actuary elevates to unforeseen heights with the right tools. Actuaries today find themselves at the disposal of a seamlessly endless collection of technical solutions that can enhance their performance in any actuarial task.

This essay will observe how deep learning, specifically image recognition, contributes to the efficiency of actuaries and reforming the industry.

Deep Learning and Image Recognition

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.¹ Such technology allows the imitation of a human mind when processing data on a large scale and has exploded in the past decade.

At the same time, image recognition, long began its development before actuarial technologies were relevant, is a deep-learning technology that allows machines to recognize and identify patterns in images, a form of unstructured data. Whether contracts or IDs, experts believe that any image can be converted into structured data for processing with enough model training. Insurance companies are now beginning to realize the importance of self-service, which eliminates repetitive documentative tasks. With numerous tech giants, namely Microsoft, and specialized companies such as OpenCV, this technology is becoming more applicable to different industries. Some applications of image recognition in insurance include:

Identification: The use of facial recognition to verify the identity of the policyholder.

Documentation: Automated documentation of medical records, police reports, and other legal documents through optical character recognition (OCR).

Personalized Policy: From their online presence in social media and other public domains, the insured's habits can now be derived. Allows for more accurate data analysis,

For a better understanding of its impact, an interview with an underwriting and claims department manager from Sino Life Insurance provided the following numbers: In 2020, 96% of

¹ Brownlee, J. (2019, December 19). What is Deep Learning? Machine Learning Mastery. <https://machinelearningmastery.com/what-is-deep-learning/>

underwriting documentation is cloud-based, 94% customer retention rate, and 60% claim settlement processed online. As this was a Chinese based insurance company, the pandemic encouraged these technological changes to emerge rapidly. ²

As physical interaction becomes improbable, OCR and facial recognition allow for the remote collection of data. The applicant can request a claim by uploading medical documents and verifying identity through facial recognition. Sino Life adopted APIs from Tencent into its claims platform that achieved an impressive 96.68% accuracy rate for facial recognition. This allowed for the reliable processing of smaller medical claims below 3000 RMB to be automatically documented and ready for inspection, which profoundly impacted their productivity.

However, we are only beginning to see the true potential of image recognition technology.

Unstructured Data

As data, the blood of actuarial science, is often what sets competitors apart in this day and age, actuarial technologies are usually oriented towards the efficiencies of data processing. One of the most promising aspects of image recognition lies in its ability to convert the unstructured data that is mostly undeveloped due to its lack of recognizable pattern for traditional AIs. In a report from Accenture, they claim that “the larger part of enterprise data, nearly 80 percent, is unstructured and has been much less accessible.”³ This accounts for emails, legal reports, etc. that are often avoided by companies due to its difficulty of organization. The report is only accounting for corporate data, and it becomes fascinating considering what lies beyond.

From traffic surveillance to satellite images, unstructured data exists in almost every aspect of a person’s daily routine, waiting to be discovered. The complexity behind insurance as it is ultimately evaluating human behavior will benefit from these data significantly.

For instance, traffic surveillance provides insight into an insured’s driving habits, and insurance companies can personalize their policy accordingly; thus, achieving first-degree price discrimination that maximizes profit. Ideally, the automated processing of unstructured data from an insured will open new data dimensions to insurance companies. ⁴

² (Xuwei, personal communication, August 2, 2020)

³ Accenture. (2020, March). Top Natural Language Processing Applications in Business. <https://www.accenture.com/acnmedia/PDF-106/Accenture-Unlocking-Value-Unstructured-Data.pdf#zoom=50>

⁴ Shang, K. (2018, June). Applying Image Recognition to Insurance. Society of Actuaries. <https://www.soa.org/globalassets/assets/files/resources/research-report/2018/applying-image-recognition.pdf>

Ethical Considerations and Challenges

That said, data collection has become quite sensitive in the past decade as consumers realize the depth of mass data collection. While this may be acceptable to consumers when browsing the internet, the opinion on purchasing insurance products becomes complicated as it could drive up rates.

Converting unstructured image data could lead to a gray zone where some information can be extracted, while others must be entirely avoided. For example, location data is highly sensitive and illegal for transfer;⁵ thus, the aforementioned traffic surveillance may be unethical as well, for it reveals personal data protected by law.

Technological difficulties are also limiting the adaptation of image recognition into insurance companies. Although speculations are promising, image recognition in insurance remains at an early stage with only OCR facial recognition accepted in insurance companies, primarily due to deficient demand across industries and underdevelopment.

A commonality between OCR and facial recognition is their reliability that is absent in the majority image cognition technologies. While the leading facial recognition solutions can achieve a 99.92% accuracy,⁶ commercial image object identification remains unreliable for business use.⁷ Neural networks have yet to reach acceptable progress for more complex pattern detection and would most likely require years of deep learning before entering commercial use.

Modern Actuaries and the Future

Actuaries nowadays are facing a technological revolution in their work as the market for actuarial tech matures. Naturally, they should begin experimenting with technical solutions that further their performance and suggest applicable technologies to their employers.

Deep learning and image recognition, while eliminating repetitive tasks, are now discovering new dimensions of data for actuarial processing. In the future, perhaps the work of an actuary will be shifted towards determining how to utilize new data for analysis. As deep

⁵ Matsakis, L. (2019, February 19). Personal Data Collection: The Complete WIRED Guide. Wired. <https://www.wired.com/story/wired-guide-personal-data-collection/>

⁶ National Institute of Standards and Technology. (2020, March). Face Recognition Vendor Test (FRVT) Part 2: Identification (NISTIR 8271). https://pages.nist.gov/frvt/reports/1N/frvt_1N_report.pdf

⁷ Enge, E. (2019, July 30). Image Recognition Accuracy Study. Perficient, Inc. <https://www.perficient.com/insights/research-hub/image-recognition-accuracy-study>

learning technology isn't limited to image recognition, automation of crucial tasks opens up an actuary's work to data explorations.

The future of data is now truly endless with possibilities.