Drone-racing champions outpaced by AI

de Croon, Guido C.H.E.

DOI
10.1038/d41586-023-02506-8

Publication date
2023

Document Version
Final published version

Published in
Nature

researched for close to 50 years. Dozens of studies have tried to understand both the nutrient 'accounting' – which explains how corals thrive in a desert-like environment – and the biology of how nutrients are acquired and moved between the partners[3]. Despite these decades of study, pieces of the puzzle have been described but a full picture that explains the success of corals in a nutrient-deficient environment has remained elusive.

Wiedenmann and colleagues grew nine coral species of interest in seawater tanks for months during which all nutrients flowing into the system were precisely controlled. The authors found that corals incubated in water containing inorganic nitrogen and phosphorus at the low concentrations typical of healthy reefs increased in size and the algal population increased as well, to keep up with coral growth. In these conditions, the partnership was acting together as an autotroph (generating its own food).

By contrast, when corals were kept in water virtually free of nutrients, after about two months, coral growth stagnated and remarkably, the algal population plummeted, causing the corals to have a bleached appearance because algae were lost. Wiedenmann and colleagues found that almost no algae were expelled into the water. Where were they going?

The authors looked for clues by tracing labelled compounds and they found that labelled nitrate was taken up by algae (the coral host lacks the machinery to take up and use nitrate). Labelled nitrate was also found in high concentrations in host tissues, indicating that nitrogen-containing compounds are transferred from the algae to their host. But how might this happen? The authors hypothesized that the algae were being consumed as food, farmed to satisfy the host's need for inorganic nitrogen and phosphorus.

To gather more evidence, Wiedenmann *et al.* examined the size of the algal population in healthy, growing corals. The authors calculated the growth rate of the algal population by counting the number of dividing cells, a measure called the mitotic index. Interestingly, when they modelled the population growth rate on the basis of the mitotic index, the expected rate was much higher than the measured rate from their experiments, obtained by counting algae over time. The authors deduced that these excess algae were being digested by the coral host.

Wiedenmann and colleagues also conducted an experiment to look for evidence of nutritional contributors to the differential coral growth in natural environments. They found that corals growing near dense seabird populations that produce high amounts of nutrient-rich guano droppings took up more inorganic nitrogen and grew faster than did corals that weren't near seabird colonies.

Nutritional flexibility of the coral–algal symbiosis might be the linchpin of the dominance of coral reefs in ocean deserts over the past 240 million years[4]. Corals can do it all. Most species can feed on other organisms and use their stinging tentacles to trap microscopic prey (zooplankton). Those corals in symbiosis with algae act as autotrophs, making sugars with energy from the Sun and recycling nitrogen and phosphorus, as well as acquiring them from the seawater. And in a pinch, when prey are not available to satisfy the corals' nutritional needs, such hosts might farm their algal population to fill the void. This scenario has evolutionary implications, given that the success of nutritional flexibility involving both partners provides evidence for the idea that the symbiotic partnership is the unit that undergoes natural selection[5], perhaps in addition to natural selection of the partners independently. The whole is therefore more than the sum of its parts.

Wiedenmann and colleagues provide strong evidence of algal farming, but there is still work to be done to definitively prove this phenomenon. Perhaps the direct evidence needed is to use an imaging method to show corals in the act of algal digestion. One study has captured images of this phenomenon for a single coral species[6]. The fact that this has not been commonly reported might mean that others have tried without success to convincingly document such a process. The authors of that imaging study[6] used electron microscopy for their work, a high-resolution but time-intensive and non-quantitative approach. Modern, state-of-the-art methods such as flow cytometry analysis of dissociated coral cells could enable high-throughput, quantitative and systematic sampling of coral tissue to look for evidence of algal degradation[7].

Finally, studies of corals in our time of extreme threats to the health of these reefs should be viewed through a lens of developing solutions to help reefs survive into the next century[8]. Knowledge of nutritional flexibility and algal farming could be used to aid understanding of the nutritional value of different algal species. It could also help to incorporate nutritional differences into strategies of breeding corals and symbionts that are more resilient to reef perturbation from temperature increases and nutrient stress. We need to marshal all available knowledge to help coral reefs remain the oases in desert oceans that are so crucial for the myriad services they provide to the planet and the people living nearby.

Virginia M. Weis is in the Department of Integrative Biology, Oregon State University, Corvallis, Oregon 97331, USA.
e-mail: virginia.weis@oregonstate.edu

1. Darwin, C. *The Structure and Distribution of Coral Reefs* (Smith Elder, 1842).
2. Wiedenmann, J. *et al. Nature* **620**, 1018–1024 (2023).
3. Morris, L. A., Voolstra, C. R., Quigley, K. M., Bourne, D. G. & Bay, L. K. *Trends Microbiol.* **27**, 678–689 (2019).
4. Stanley, G. D. *Earth Sci. Rev.* **60**, 195–225 (2003).
5. Bordenstein, S. R. & Theis, K. R. *PLoS Biol.* **13**, e1002226 (2015).
6. Titlyanov, E. A. *et al. Mar. Ecol. Prog. Ser.* **139**, 167–178 (1996).
7. Rosental, B., Kozhekbaeva, Z., Fernhoff, N., Tsai, J. M. & Traylor-Knowles, N. *BMC Cell Biol.* **18**, 30 (2017).
8. Kleypas, J. *et al. Biol. Conserv.* **257**, 109107 (2021).

The author declares no competing interests.
This article was published online on 23 August 2023.

Engineering

Drone-racing champions outpaced by AI

Guido C. H. E. de Croon

An autonomous drone has competed against human drone-racing champions – and won. The victory can be attributed to savvy engineering and a type of artificial intelligence that learns mostly through trial and error. **See p.982**

Artificial intelligence (AI) has already surpassed the performance of human champions in games such as chess[1], Go[2] and the car-racing video game *Gran Turismo*[3]. However, these achievements all took place in virtual environments. On page 982, Kaufmann *et al.*[4] make the leap to the real world with Swift – an autonomous AI-based drone system that can defeat humans in the sport of drone racing. Swift took on three human adversaries, all of whom are drone-racing champions, and clocked the fastest time on the racetrack.

The vehicles used in drone racing are usually controlled by human pilots who wear headsets that give them a 'first person' view through a camera attached to the drone. These pilots manoeuvre the drones deftly through a series of gates at speeds of 100 kilometres per hour or more (Fig. 1). The gates are positioned to make the tracks difficult, pushing the drones to undergo accelerations several times that of gravity.

Drone racing is highly exacting for experienced human drone pilots, but it poses even more challenges for AI. In a virtual environment, resources are practically endless; moving to the real world means having to work with limited resources. This is especially true of drones, because the sensors and computing that replace the human pilot must be lifted into the air. Kaufmann and colleagues' solution was to minimize sensing equipment: their drone's sensor suite comprises a camera with a wide field of view, a small binocular vision system and a tiny inertial measurement unit — the device that measures acceleration and rate of rotation in a smart phone.

The real world is also much more unpredictable than a virtual one. Whereas a simulated race car could cruise perfectly along the trajectory that is programmed for it, a single remote-control command issued to a drone can have many unexpected effects. Indeed, although the flight behaviours of the drones typically used in racing have been modelled extensively[5], predictions still fall short of reality, especially at high speeds, during agile manoeuvres or in the presence of external disturbances[6]. These real-world effects are difficult to predict, and using the resulting imperfect models is especially problematic for drones that are trained with AI, because such approaches typically rely on learning in a simulated world.

The standard way of training an AI robot is called end-to-end learning. For drone racing, this approach involves learning how to map images to commands that specify the speed of the drone's propellers. But the set of mappings learnt in a virtual environment often fails to transfer to the real world because of the differences between the simulated and real systems. This problem is called the reality gap, and it constitutes one of the main challenges in AI for robotics.

Kaufmann and colleagues' work is a great example of how roboticists are overcoming the reality gap[7]. Swift is trained using a judicious combination of AI-learning techniques and conventional engineering algorithms. The approach involves first processing the images that the drone obtains with the camera, using an artificial neural network that can detect gate corners — a task at which AI excels, given sufficient training data, curated by humans. The drone's speed is then determined using proprietary software that comes with the binocular vision system.

These two sets of information (the vehicle's speed and the gate locations) are then incorporated with data from the inertial measurement unit using conventional algorithms that estimate the vehicle's state — the same algorithms that feature in car navigation systems. The state variables here are the drone's position, velocity and attitude (its orientation with respect to gravity).

Swift's true innovation, however, is another artificial neural network that maps the drone's state to commands that modify its thrust and rate of rotation. This network uses reinforcement learning[8], a technique that optimizes a reward received from the environment by means of trial and error in simulation. Applying reinforcement learning in this part of the algorithmic pipeline, instead of end-to-end learning, allowed the authors to cross the reality gap using a concept known as abstraction[9].

Because the state variables encode a higher level of abstraction than the raw images, the reinforcement-learning simulator does not need to render a rich visual environment. Given good state estimation, this reduces the differences between the simulated and real systems considerably. Moreover, it leads to much faster simulations, enabling the system to learn in around 50 minutes.

The actions taken by the network (that is, applying the desired thrust and rotation rates) are also at a higher abstraction level than the commands used to change the rotation speed of its propellers. This means that the AI system can execute these actions reliably using an onboard controller, enabling the drone to cope with disturbances and effects that are not included in the simulation model. The small differences that remain between simulation and reality are learnt by a neural network to improve the simulation and refine the system's strategy.

Clearly, the impact of Kaufmann and colleagues' achievement extends well beyond drone racing. The obvious, if controversial, possibility is that this technology could find a military application — a central preoccupation of many roboticists working on AI-based drones. However, the results have a much broader range of applications. A decade ago, most autonomous drones were excruciatingly slow, getting very little range out of their batteries. Although extremely fast flight will not be needed for most real-world uses, the techniques developed by Kaufman *et al.* will allow for smoother, faster and longer-range missions than those possible with existing drones. Moreover, they will help all robots, whether they are used for driving, cleaning or inspecting, to get more out of their limited onboard resources.

But to realize this potential, further developments will be necessary. The authors' tests took place in a controlled indoor environment, but real drone races are held in varying environments, both indoors and outdoors. To beat human pilots in any racing environment, the drone will have to deal with external disturbances such as the wind, as well as with changing light conditions, gates that are less clearly defined, other racing drones and many other factors — all of which pose sizeable challenges to existing AI techniques.

Given that drones acquire sensing information more rapidly than do human pilots, who rely on delayed images, drones will no doubt eventually beat humans under these difficult conditions as well. The future could well turn, then, to ever-faster drone races that

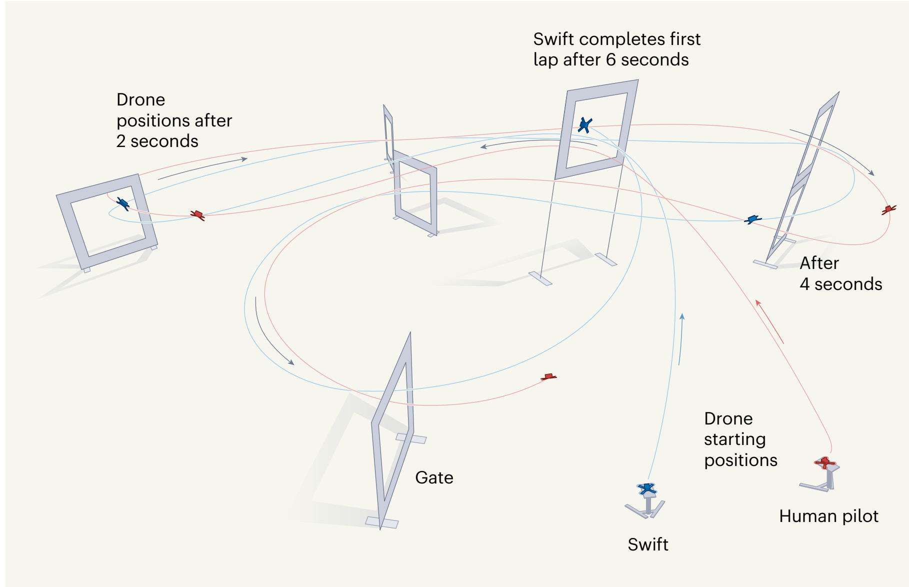

Figure 1 | A course for drone racing. Kaufmann *et al.*[4] designed an autonomous AI-based drone system called Swift, and demonstrated that it could outpace humans who are champions in the sport of drone racing. The authors set a course comprising seven gates, which had to be passed through in order. The drones had to complete three laps — the first of which is shown here, for Swift and one human competitor. Swift beat three champion-level human drone-racing pilots and established the fastest time on the track. (Adapted from Fig. 1a of ref. 4.)

pit autonomous drones against each other[10] — a development that will keep pushing the boundaries of this widely relevant technology.

Guido C. H. E. de Croon is in the Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS Delft, the Netherlands.
e-mail: g.c.h.e.decroon@tudelft.nl

1. Campbell, M., Hoane, A. J. Jr & Hsu, F. H. *Artif. Intell.* **134**, 57–83 (2002).
2. Silver, D. *et al. Nature* **550**, 354–359 (2017).
3. Wurman, P. R. *et al. Nature* **602**, 223–228 (2022).
4. Kaufmann, E. *et al. Nature* **620**, 982–987 (2023).
5. Beard, R. W. *Quadrotor Dynamics and Control* (Brigham Young Univ., 2008).
6. O'Connell, M. *et al. Sci. Robot.* **7**, eabm6597 (2022).
7. Salvato, E., Fenu, G., Medvet, E. & Pellegrino, F. A. *IEEE Access* **9**, 153171–153187 (2021).
8. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 2018).
9. Scheper, K. Y. W. & de Croon, G. C. H. E. In *Proc. 14th International Conference on Simulation of Adaptive Behavior* (eds Tuci, E., Giagkos, A., Wilson, M. & Hallam, J.) 280–292 (Springer, 2016).
10. De Wagter, C., Paredes-Vallés, F., Sheth, N. & de Croon, G. C. H. E. *Field Robot.* **2**, 1263–1290 (2022).

The author declares no competing interests.

Clinical neuroscience

Speech-enabling brain implants pass milestones

Nick F. Ramsey & Nathan E. Crone

Two brain–computer interfaces have been developed that bring unprecedented capabilities for translating brain signals into sentences — at speeds close to that of normal speech, and with vocabularies exceeding 1,000 words. **See p.1031 & p.1037**

There is an urgent need to help people with neurological conditions that deprive them of the universal human need to communicate. Two articles published in *Nature* demonstrate that individuals who are unable to speak as a result of severe paralysis could potentially use implantable brain–computer interfaces (BCIs) to communicate at rates much greater than those typically achievable with alternative communication options. Willett *et al.*[1] (page 1031) report a device that records brain activity using electrodes that penetrate the brain's cortex, whereas Metzger and colleagues' device[2] (page 1037) uses electrodes placed on the cortical surface. These studies signal a turning point in the development of BCI technology that aims to restore communication for people with severe paralysis.

Various neurological disorders paralyse muscles crucial to speech and limb function while sparing cognitive functions, potentially resulting in locked-in syndrome — in which individuals can no longer initiate communication and can respond to queries only with eye blinks or minimal movements. A diverse range of systems, known as alternative and augmentative communication technologies, are available to help people with locked-in syndrome to communicate, but these require effort and are much slower (achieving, typically, just a few words per minute) than normal speech (about 150 words per minute). BCIs have the potential to solve these problems.

The first demonstration that a subject could be trained to increase the activity of single neurons, and thereby to exert a wilful action, was published in 1969, for a rhesus macaque (*Macaca mulatta*)[3]. Experiments in humans began[4] in the late 1990s, when an electrode was connected to neurons in a person with locked-in syndrome caused by motor neuron disease (amyotrophic lateral sclerosis, or ALS), a neurodegenerative disease. This was followed in 2006 by a study[5] in which arrays of millimetre-scale electrodes (known as microelectrodes) were implanted into the brain of a person with a spinal cord injury. This microelectrode array (MEA) recorded the activity of several hundred neurons in the motor cortex, the brain region responsible for the control of voluntary movements, and thereby controlled a robotic arm[5]. MEAs have since been used to enable communication, for instance by decoding handwriting attempts[6].

The complementary technique of electroencephalography (EEG) — in which electrodes are placed along the scalp to record electrical activity in the brain — has been used since 1999 (ref. 7) to help people with paralysis to communicate by controlling custom spelling software[8]. Around the same time, it was discovered that small disc-shaped electrodes (2–3 millimetres in diameter) placed on the surface of the brain could acquire much higher-quality signals than could be obtained using scalp electrodes[9]. This method for recording brain activity is known as electrocorticography (ECoG).

In the early 2000s, ECoG electrodes were used in people undergoing surgery for drug-resistant epilepsy, to record brain signals associated with speech and body movements[10]. This eventually led to the development of the first fully embedded ECoG device, which enabled a person with locked-in syndrome to use a typing program at home[11]. To date, about 50 people with varying degrees of paralysis have been implanted with BCIs for communication, most of whom use MEAs.

Metzger *et al.* now present findings involving a paralysed participant who, 17 years before she enlisted for the study, experienced a brainstem stroke that made her speech unintelligible. The authors' BCI system incorporates a silicon sheet embedded with 253 ECoG electrodes, each of which record the average activity of many thousands of neurons (Fig. 1a). The device was surgically implanted over the left 'face area' of the sensorimotor cortex — the part of the brain that serves oral and facial muscles, including the vocal tract. The study builds on previous reports of ECoG recordings, including a similar BCI that was implanted in another person who had had a brainstem stroke[12].

Brain-to-text decoding was achieved by the combination of two systems: a recurrent neural network (RNN, a type of artificial neural network), which ran algorithms that decipher brain activity associated with movements of articulators (parts of the vocal tract); followed by a language model that composed sentences at a rate of 78 words per minute (albeit with a 25.5% word error rate) from a set of 1,024 words. Alternatively, brain signals were translated directly to synthesized speech, at a word error rate of 54.4% for the 1,024-word vocabulary; the error rate decreased for smaller vocabularies (8.2% for a 119-word vocabulary). The BCI also decoded attempted facial expressions, which it reproduced using a digital avatar, thereby providing visual feedback to the text or speech that greatly enriches the participant's ability to communicate. Overall, the device offers substantial improvements in vocabulary size, speed of communication and versatility of speech decoding compared with previously reported ECoG BCIs.

Willett *et al.*[1] used two MEAs (containing a total of 128 electrodes) to record from small patches of the left sensorimotor face area in a participant who was unable to speak intelligibly owing to ALS (Fig. 1b). As in Metzger and colleagues' device, RNNs and language models were used to translate brain signals into text and were trained and tested on vocabularies of different sizes. Using the device, the participant was able to communicate at an average rate of 62 words per minute, with a word error rate of 23.8% for a 125,000-word vocabulary and 9.1% for a 50-word vocabulary.

The RNN was trained using recordings of neural activity collected when the participant attempted to speak 260–480 sentences presented on a monitor — the overall process