CERBERUS in the DARPA Subterranean Challenge

Review Article

Author(s):
Tranzatto, Marco; Miki, Takahiro; Dharmadhikari, Mihir; Bernreiter, Lukas; Kulkarni, Mihir; Mascarich, Frank; Andersson, Olov; Khattak, Shehryar Masaud Khan; Hutter, Marco; Siegwart, Roland; Alexis, Kostas

Publication date:
2022-05-25

Permanent link:
https://doi.org/10.3929/ethz-b-000551504

Originally published in:
Science Robotics 7(66), https://doi.org/10.1126/scirobotics.abp9742

CERBERUS in the DARPA Subterranean Challenge *

Marco Tranzatto[1*], Takahiro Miki[1], Mihir Dharmadhikari[2], Lukas Bernreiter[4], Mihir Kulkarni[2], Frank Mascarich[3], Olov Andersson[4], Shehryar Khattak[1], Marco Hutter[1], Roland Siegwart[4], Kostas Alexis[2]

[1]Robotic Systems Lab, ETH Zurich

[2]Autonomous Robots Lab, Norwegian University of Science and Technology

[3]Autonomous Robots Lab, University of Nevada, Reno

[4]Autonomous Systems Lab, ETH Zurich

*To whom correspondence should be addressed; E-mail: `marcot@ethz.ch`

This article presents the core technologies and deployment strategies of Team CERBERUS that enabled our winning run in the DARPA Subterranean Challenge finals. CERBERUS is a robotic system-of-systems involving walking and flying robots presenting resilient autonomy, as well as mapping, and navigation capabilities to explore complex underground environments.

This article details the winning performance of Team CERBERUS in the DARPA Subterranean Challenge Final Event.

Introduction

The DARPA Subterranean (SubT) Challenge was a three-year, $82 million (*1*) robotics competition organized and coordinated by the US Defense Advanced Research Projects Agency

*This manuscript has been accepted for publication in Science Robotics. This version has not undergone final editing. Please refer to the complete version of record at `https://www.science.org/doi/10.1126/scirobotics.abp9742`.

(DARPA). The SubT Challenge pushed roboticists to develop innovative technologies to support operations in complex underground settings, presenting notable challenges for military and civilian first responders. The subterranean presents extreme environmental diversity, ranging from tunnels spanning over multiple kilometers to complex urban multi-level settings, from narrow passages and vertical shafts connecting cave networks to vast chambers, and from areas that are seamless to traverse to grueling sections. Hence, operating in these environments is difficult and can often be dangerous for humans. Therefore, robotic systems can be a valid alternative to provide rapid situational awareness to a small team of operators prior to possible entry in such unknown and dynamic subterranean areas. This competition brought together over 300 competitors from 20 different teams, composed of individuals, startups, universities, and large companies that spanned 11 countries and 20 universities.

The SubT Challenge was organized into two competition tracks: Systems and Virtual. Systems track teams deployed their hardware and software solutions to compete in physical subterranean courses. In contrast, Virtual track teams developed software-only solutions evaluated in simulation. Synergy among the two tracks was facilitated especially by the release of simulation models of the robots teams deployed in the Systems track. The underground scenarios for the Systems track included human-made tunnel networks extending for several kilometers and presenting vertical openings ("Tunnel Circuit" - August 2019), multi-level urban underground structures with complex layouts ("Urban Circuit" - February 2020), and natural cave environments with complex morphology and constrained passages ("Cave Circuit" - August 2020, canceled event). Finally, a final, comprehensive, benchmark mission combined all the challenges of the previous circuits ("Final Event" - September 2021).

Competing teams were tasked with exploring and mapping unknown underground areas, with the additional requirement that only one team member, the Human Supervisor, was allowed to manage and interact with the deployed robots. Teams were scored based on their ability to localize specified objects, called “artifacts”, representing features or items of interest applicable to subterranean settings. Artifacts for the Prize Run of the Final Event included human (mannequin) survivors, cellphones, backpacks, drills, fire extinguishers, vents, gas-filled rooms, helmets, ropes, and a custom SubT cube. Each team had 60 minutes to find as many as possible of the 40 artifacts distributed along the course, while having a fixed number of available reports (to discourage false positives). A point was earned for each report that correctly identified an artifact’s class, and its position within five meters of the object’s ground-truth location.

DARPA initiated the SubT Challenge to stimulate development in four technical areas: Autonomy, Perception, Networking, and Mobility. Robotic systems were asked to show a high level of resilient autonomy, proving to be able to map, navigate, and search in complex and dynamic environments without substantial human interventions (Autonomy). Additionally, they had to cope with hard navigation challenges, operating under varying and degraded conditions, involving dust, fog, mist, water, smoke, darkness, obscured views, and/or scattered environments (Perception). Moreover, the robots had to send information to the Human Supervisor while dealing with limited line-of-sight, the effects of varying geological conditions, and wireless signal propagation challenges in subterranean environments (Networking). Finally, the systems were tasked to navigate mobility-stressing and dynamic terrain, including narrow passages, sharp turns, large drops/climbs, inclines, steps, falling debris, mud, sand, and water (Mobility).

The Prize Round of the Final Event took place at the Louisville Mega Cavern in Kentucky, in September 2021 and involved eight teams for the Systems track and nine teams for the Virtual track. For the Systems track, other than Team CERBERUS, the following teams competed: “CSIRO Data61” (deploying tracked vehicles, legged robots, and drones), “MARBLE” (leveraging legged and wheeled robots), “Explorer” (relying primarily on custom-built wheeled and flying robots), "CoSTAR" (using legged, wheeled, and drones), "CTU-CRAS-NORLAB" (deploying legged, tracked, and flying robots), "Coordinated Robotics" (using wheeled and flying robots) and "Robotika" (using wheeled robots). In the leaderboard of the Systems track's prize run, both teams "CERBERUS" and "CSIRO Data61" scored 23 points, followed by teams "MARBLE" (18), "Explorer" (17), "CoSTAR" (13), "CTU-CRAS-NORLAB" (7), "Coordinated Robotics" and "Robotika" (both with 2 points). As both CERBERUS and CSIRO Data61 were tied with 23 points each, the tiebreaker rule had to be invoked: "the team that identified its last artifact earliest would win." CSIRO Data61 reported the final point with under 30 seconds remaining, while CERBERUS reported its last artifact with more than 1 minute left thus breaking the tie and winning the DARPA Subterranean Challenge, as well as a $2 million prize reward.

Development and Deployment Strategy

Team Introduction

Team CERBERUS (*2*) represented a collaborative system-of-systems (Fig. 1) involving walking and flying robots equipped with multi-modal perception capabilities, navigation and mapping autonomy, and self-deployed network communication modules. This enabled the reliable access, navigation, exploration, mapping, and object search in complex, sensing-degraded, dynamic, and rough subterranean environments. A tethered roving platform assisted the team and acted as a communication hub. As developed by the end of the SubT Challenge, CERBERUS involved multiple ANYmal C legged robots that offered the necessary endurance and the capacity to traverse challenging terrain. Flying robots, small-sized collision-tolerant platforms and larger multirotor systems, accompanied the walking platforms, thus permitting the access and exploration of environments with a substantial height such as mine stopes, shafts, and large cave chambers. All robots implemented resilient multi-modal sensor fusion exploiting LiDARs, as well as visual or thermal cameras (optional), that could collectively penetrate and localize in environments that span over several kilometers, involve self-similar geometries, and present sensor-degradation challenges including dense obscurants such as smoke and dust. A unified exploration path planning strategy allowed the CERBERUS system-of-systems to explore diverse subterranean environments autonomously. To gather the necessary expertise, Team CERBERUS leveraged international partnerships involving the University of Nevada Reno, ETH Zurich, the Norwegian University of Science and Technology (NTNU), University of California Berkeley, University of Oxford, Flyability and Sierra Nevada Corporation.

Figure 1: Instances of field deployments of a subset of the CERBERUS system-of-systems robots in the DARPA Subterranean Challenge Final Event environment and other testing sites. The ANYmal C SubT quadruped robot exploring and looking for artifacts in urban settings (A), tunnel environments (B), and cave networks (C). The tethered roving platform driving in a cave network to act as communication hub for the other robots (D). Medium-sized aerial scout (E) and RMF-Owl (F) exploring sections of underground mines that may involve high vantage points.

Robots and Sensors

CERBERUS's legged agents were based on the "ANYmal C100" series quadruped by ANYbotics, customized to create the ANYmal "SubT" version. This version was water- and dust-proof (IP67) and could operate in humid and dusty conditions. It weighed 55 kilograms including the payload and could operate for 80 minutes while continuously walking. There were two types of ANYmal C SubT robots: the "Explorer" and the "Carrier", which differed in their sensor configuration and mission role. The Explorer's goal was to proceed deep in the subterranean environments while detecting as many artifacts as possible. The Carrier robots had fewer sensors and supported underground operations by carrying and deploying communication-extender modules on the ground to increase the wireless network range. Both robot types were equipped with a LiDAR sensor used for Simultaneous Localization And Mapping (SLAM) and an "Alphasense Core," a visual-inertial sensor by Sevensense, which encapsulated three monochrome and four color cameras used both for SLAM and to detect artifacts. Moreover, the Explorer robots were equipped with two additional LiDAR sensors for improved terrain mapping and a pan-tilt unit head featuring a zoom-enabled camera, a thermal camera, a microphone, and a spotlight.

CERBERUS utilized a team of "Aerial Scouts" to provide fast exploration, especially in the areas not accessible by the ground robots. There were three main categories of these scouts, namely medium-sized multirotors (*3*), small-sized collision-tolerant flying robots ("Gagarin" and "RMF-Owl"), and a single large-sized tricopter ("Kolibri"). The RMF-Owl (*4*) and the Gagarin (*3*) platforms integrated LiDAR and vision and were two types of collision-tolerant aerial vehicles featuring a protective outer cage. Gagarin was developed around the "Elios" collision tolerant platform built by Flyability. The medium-sized scouts, built around the "M100 frame" by DJI, integrated a LiDAR, a color camera, as well as a thermal camera and could penetrate dusty and smoke-filled environments. The large-sized Kolibri, based on a tricopter platform from Voliro, was intended for long-term missions (over 20 minutes endurance) and vertical exploration (by being able to pitch independently of its translation).

An additional wheeled ground vehicle extended the CERBERUS communication network. This rover featured a high-gain directional antenna and a 300 meter-long optical fiber reel connected to the Base Station with the capability of automatic unrolling while the rover moved.

Modes of Operation

The CERBERUS system-of-systems functioned in a "supervised autonomy" mode by exploiting several degrees of robotic autonomy for exploration and search of underground environments, combined with human reasoning to guide the deployment strategy of the robotic team. It was "supervised" in that the Human Supervisor reasoned about the environment information available at the Base Station and, based on their situational assessment, made high-level decisions and could assign subtasks to each robot towards one collective exploration strategy. Each system then undertook the execution of its task autonomously. First, a robot could be tasked to explore a certain unknown area while ensuring it remained within communication range. If it went beyond the communication range after a certain step, it would backtrack to a point with connectivity. Second, a robot could also be tasked to explore unknown regions away from communication range, given a set time budget after which it would return back to connection range. A third mode allowed the Supervisor to command the robot to reach a frontier close to unmapped areas and then trigger autonomous exploration. Finally, as a last resort, the Supervisor could provide a waypoint in the map and request the robot to find an admissible path to reach it. Overall, in the CERBERUS' vision, highly autonomous robots became part of an exploration strategy driven by human reasoning in the described supervised autonomy fashion.

Overview of the Result

Fig. 2 shows an overview of the winning prize run, including the collectively explored map and the trajectories (A) of the four deployed ANYmal C robots, some of the challenges faced during the mission (B to E), and a selection of our agents' exploration, perception, and autonomy capabilities (F and G). Overall, the robots walked for more than 1.7 kilometers and successfully identified and reported 23 artifacts out of 40 that were distributed along the course.

Technologies for Underground Autonomous Operations

Resilient Autonomy

Search for objects of interest in unknown and communication-limited subterranean environments necessitates that robotic systems are capable of highly autonomous exploration with little-to-no human supervision. Responding to this challenge, CERBERUS implemented a resilient autonomy paradigm that uniformly allowed both legged and flying robots to access, explore, map, and search diverse underground environments. At the core of this autonomy solution was the Graph-based exploration path planner (GBPlanner) (*5, 6*). Exploiting a volumetric representation of the environment (*7*), GBPlanner employed a bifurcated architecture to deliver efficient "local exploration" alongside "global autonomy" over kilometer-sized maps. The local exploration stage was normally triggered and started by querying the robot location, the currently explored map (created using the sensor measurements and robot poses), and some defined bounds of the exploration space. First, a random graph within a local bounding box was sampled. The method then identified the path that the robot could take within the explored space, such that its range sensors could explore more unknown space efficiently. The algorithm allowed for true 3D exploration, while simultaneously considering the admissibility limitations of slopes and other terrain anomalies for ground systems. Due to the complex topologies and often highly-inclined terrain presented by the subterranean environments, the local exploration

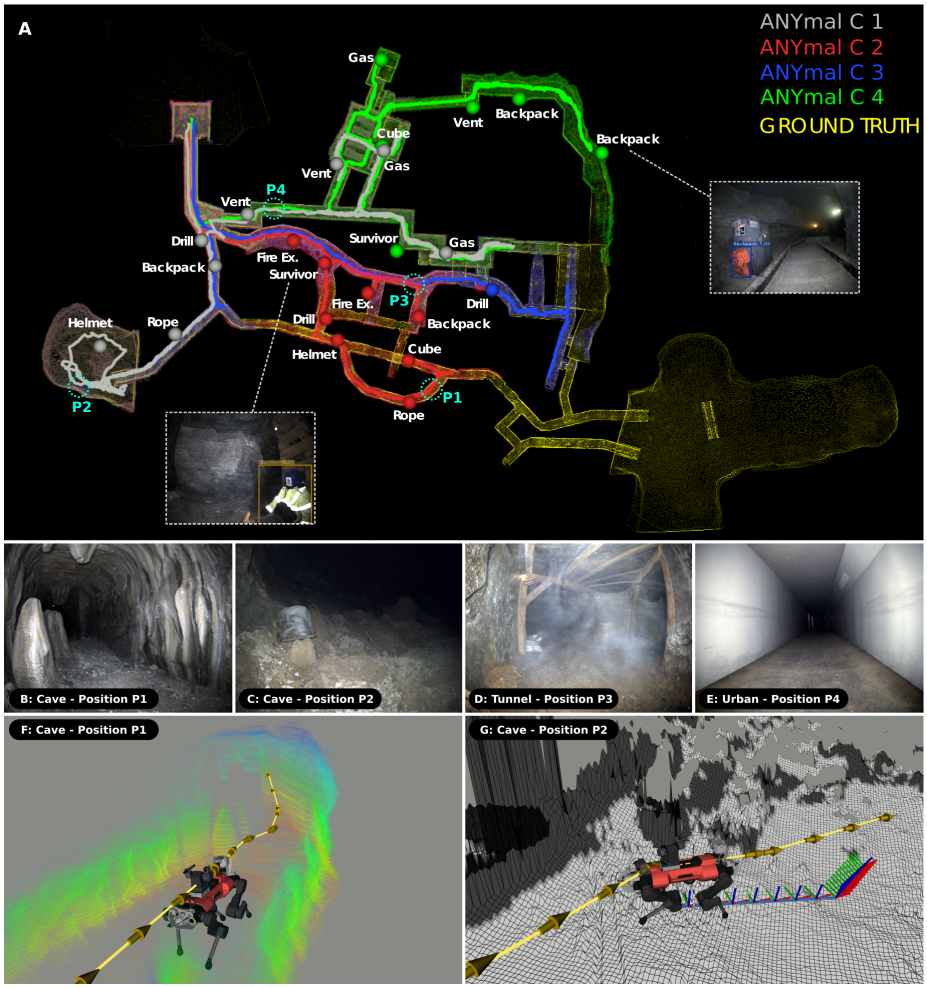

Figure 2: Overview of the winning prize run of Team CERBERUS in the DARPA Subterranean Challenge Final Event. The collectively explored map (A), the trajectories of the four deployed ANYmal C SubT robots, and the 23 successfully reported artifacts with their locations (with examples of detected "survivor" and "backpack" artifacts). The collective map is composed of the maps created by each robot (colorized based on the agent) and overlapped on the ground truth map (in yellow) provided by DARPA. The robots faced multiple challenges during their mission, as shown in the images collected with the onboard cameras: narrow passages (B), rough terrains (C), sensor degraded tunnels filled with smoke (D), and self-similar environments with a lack of visual texture (E). Each robot generated exploration paths (in yellow) to explore unknown underground environments and proceed further to search for artifacts (F). Each exploration path (in yellow) was adjusted by the reachability-based navigation planner to create a traversable path (in blue with axes markers) based on the traversability proprieties encoded in an elevation map (G), where areas displayed in white were classified as steppable.

stage may not be able to find a path leading to further exploration. At this moment, the planner could — automatically or through possible human supervision — trigger its global mode. The global stage exploited an incrementally built sparse graph that spanned across the known map and stored information about the frontiers of exploration along with the paths going to them. This allowed auto-re-positioning to unexplored areas, as well as automated or human-triggered return to home or to a previous point of communication with the Base Station either due to battery limitations or when a set time allotment expired.

Robust Mobility

Given the path from the exploration planner, the ground and flying robots needed to navigate the environment to perform the exploration. The Final Event course contained various challenging obstacles for robot mobility. The tunnel area had rails or hose tubes, the stairs in the urban area posed a challenge for ground vehicles, and the cave area presented a rough and slippery surface. The ground contained gravels, rocks, and a steep and narrow passage posing navigation and locomotion challenges.

To avoid local obstacles and follow the exploration path safely, exploration waypoints provided by the exploration planner were followed by a legged robot-specific reachability-based navigation planner (*8*) that commanded the quadruped's locomotion controller. This navigation planner could locally adjust the exploration path to better cope with the properties of the surrounding terrain. It used a local GPU-derived elevation map (*9*) and a cost map that encoded traversability properties of the area.

Moreover, our learning-based perceptive locomotion controller (*10*) was employed on the ANYmal robots to track the velocity commands generated to follow a local, traversable, path. This controller was entirely trained in simulation using reinforcement learning and learned how to jointly generate commands for the robot's actuators to enable it to walk over various types of rough terrain. We applied data randomization during training to achieve a high level of robustness. Namely, we randomized external forces and torques on the main body, the initial body and joint velocities, the total mass, and the slipperiness of the terrain. The key aspect of this controller was to combine the robot's onboard multi-modal information robustly. It was critical to perceive the environment using the exteroceptive measurement from the local elevation map to handle obstacles such as stairs, however, due to the challenging navigation conditions applicable in the subterranean environment, the sensors could get degraded for multiple reasons. In fact, the course contained dust, smoke, and water puddles, which added noise to the sensor measurements and made them unreliable. We trained our controller with various exteroceptive noise sources to deal with this case. The policy learned how to balance the amount of information to use from each modality through end-to-end training (*10*). Our locomotion controller could decide whether to use exteroceptive inputs, if detected as reliable, for fast and smooth locomotion or choose to ignore them if marked as degraded. Moreover, the capability of dealing with noisy measurements also made the controller able to cope with sudden sensors malfunctioning (i.e., no more measurements available). In previous phases of the competition, we explored different foot configurations (*3*) (point-foot, flat-foot, wheels) for the ANYmal robot. We converged that a point-foot design offers the best versatility against terrain diversity and provides a comprehensive solution, compared to wheeled or tracked robots, to overcome most terrain without needing to specialize their design.

With respect to our aerial robots, our systems employed the conventional paradigm of cascade position-attitude control exploiting fixed gain structures or model predictive control strategies (*3, 4, 11*).

Resilient Localization and Mapping

For reliable robot navigation in complex subterranean environments, robust localization and mapping are among the core components of robot autonomy. These components are responsible for providing consistent and low-latency localization for real-time robot operations and producing an accurate and informative map. In response, the CERBERUS' localization and mapping solution was based on two key understandings. First, multi-modal sensing was crucial to enable reliable robot operation in the underground, by providing redundancy and access to diverse multi-modal sensor data inputs that present different failure points. Second, the developed solutions were flexible to adapt to variations in sensor payloads and compute capabilities present across the diverse ground and flying robots of CERBERUS. Building on these ideas, a two-stage localization and mapping solution was developed.

For onboard localization and mapping, each robot estimated its pose and map in real-time fully onboard using a complementary multi-modal sensor fusion approach, CompSLAM (*12*), making each robot an independent autonomous agent. CompSLAM fuses visual and thermal imagery, LiDAR depth data, inertial cues, and kinematic pose estimates in a hierarchical manner to provide redundancy against cases of sensor data degradation and estimation failures. The key idea was that for each sensing modality, or a subset, an independent robot pose estimator was operational — the output of which, subject to passing health checks, was provided as an initialization point for the next sensor pose estimator in the hierarchy. If an estimator was determined to be inconsistent, due to data- or process-level health checks failure, the pose of the previous estimator was provided directly to the next pose estimator, skipping the inconsistent estimator in-between. This design robustly ensured that a robot always had an updated pose estimate and map available.

To exploit the multi-robot constraints for the joint optimization of robot poses and to build a mission map for artifact reporting, a multi-modal multi-robot mapping solution, M3RM (*13*), was developed. This solution is a centralized approach that takes advantage of the extensive computational resources on the Base Station to improve localization and mapping at the joint mission level. When in communication range, individual robots incrementally transferred their odometry estimates and sensor data to create an inter-robot factor graph. Each robot sent visual-inertial factor graphs with subsampled and compressed LiDAR measurements to account for the network's bandwidth limits. This graph was optimized to create a globally consistent multi-robot map and report optimized artifact locations. As a fallback, the system could use the CompSLAM solutions per robot for artifact reporting, while most aerial robots only reported their onboard solution regarding their pose and map estimates. Overall, the bifurcation of localization and mapping at the robot and mission levels allowed for improved resilience by avoiding a single point of failure.

Communications

Underground settings can be particularly demanding for wireless communication as signal propagation is often unpredictable and highly influenced by the environment geometry. The autonomy capabilities of each robot are a crucial feature to ensure the exploration of subterranean environments with minimal human supervision. Still, at the same time, extended wireless network coverage is essential to swiftly send the gathered information (map, objects of interest) from the robots to the Base Station. Therefore, CERBERUS employed an ad-hoc wireless mesh network, composed of several "nodes," to cover extended distances and ensure a connection link with the robots exploring deep in the unknown. This network operated in the 5.8 gigahertz range and was composed of the legged robots and the wheeled ground vehicle, which acted as nodes. Carrier ANYmal robots could ferry and deploy additional communication-extender modules, also acting as mesh nodes, to increase the reach of the wireless mesh network incrementally upon human command. Each Carrier ferried four of these modules on its back that were released by lowering the torso height and activating an electromagnetic release mechanism. Each module was a water- and dustproof self-contained unit, including a "BreadCrumb DX2" radio by Rajant and a battery that allowed two hours of operation. The mesh network constituted the team's backbone communication infrastructure, which could also be accessed by the aerial scouts, acting as clients.

Artifact Detection

Accurate artifact detection and localization was key to scoring in the SubT challenge. To detect the visual artifacts (all except cellphones, gas and the SubT cube), Team CERBERUS used YOLOv3 (*14*) trained using a total of over $40,000$ labeled images collected in subterranean environments. Beyond collecting a large amount of data, attention was paid to gathering artifact images in the presence of obscurants, using different cameras, diverse background appearances, as well as with varying light conditions and motion blur. The trained network was deployed onboard each robot. For each artifact detected in an image, rays were cast from the pixels inside the detection bounding box into the robot's volumetric map to obtain an initial position of the artifact. Subsequent measurements collected as the robot moved allowed improving artifact location estimates. False detections were prevented using binary probability filters that updated the probability of the artifact class with each detection. The cellphone and the SubT cube were detected using Bluetooth signal strength measurements, and gas with a CO2 sensor. Given a window of measurements, the artifact location was set to the robot's location where the maximum intensity was acquired. The Human Supervisor inspected and accepted/rejected a detection prior to submitting an artifact report.

Lessons learned

The DARPA Subterranean Challenge exposed robotic technologies to diverse challenges and environments, offered multiple lessons based on real-life field-testing, and allowed researchers to draw important conclusions that influence future robotics developments. Although we developed the presented methods and systems mainly to address the challenges posed by the SubT competition, we expect that they can be adapted and deployed in different fields other than subterranean exploration (such as mapping of construction sites, exploration of new infrastructures, ship and process tank inspection, etc.). We attribute this potential for generalization to the similarity of the underlying challenges in terms of arduous terrain (e.g., in construction), perceptually-degraded conditions (e.g., in industrial assets), and the need for autonomous mapping and object recognition for a host of inspection and monitoring tasks. Below, we summarize our most pertinent lessons learned.

Legged and flying robots are the ideal combination for subterranean environments. In the beginning of the challenge, teams explored different options including wheeled, tracked, legged, and flying robots. CERBERUS specifically proposed an almost complete reliance on walking and flying systems. At the Final Event, almost all teams heavily relied on legged and flying robots. Walking robots represented the prime exploration unit offering both endurance and the ability to overcome challenging terrain in a manner superior to wheeled or tracked systems. Flying robots were the complementary system to "fill the gap" and explore highly vertical environments, and impassable terrain. However, small aerial robots suffer from limited endurance, while it is possible that the areas where such systems are mostly needed (e.g., a vertical passage) are far from the deployment point of our systems. Responding to this fact, our team aimed for the marsupial ferrying of RMF-Owl by ANYmal (*15*), although this development was finalized only after the competition. This can offer the complementary capabilities of a flying robot when and where the legged robots could truly benefit by its support deep in the underground.

Resilient perception and autonomy are the key technologies to explore and search subterranean environments. Despite historic progress in SLAM and path planning, underground environments proved to be the domain where autonomy had many reasons to break. Systems had to be robust, redundant, and resourceful in their ability to enable perception through the most degraded conditions and within their effort to explore deeper on their own. Multi-modal robotic perception fusing complementary sensors capable of penetrating diverse conditions of perceptual-degradation, and presenting different cases of ill-conditioning (e.g, LiDAR is unaffected by lack of texture which however challenges visible-light cameras), was essential in order to minimize possible breaking points in the robots' ability to localize and map. Furthermore, we have demonstrated that an exploration planning policy reflecting the key geometric challenges of underground environments could uniformly guide vastly different robotic configurations to explore and search efficiently. However, some key aspects are lacking in autonomy to unlock the full potential of the robots which include semantic understanding of the environment, a comprehensive modeling of navigation risks (e.g., the interplay between perception conditions and path planning decisions), tigher coupling of exploration planning and traversability constraints (when applicable), (multi-)robot SLAM with improved consistency against severe sensor degradation, and efficient decentralized multi-robot coordination.

The "dialectic" synergy between robotic autonomy and human input. Subterranean environments, unless instrumented a priori, prohibit the reliable self-carried high-bandwidth communication among the robots and to the Base Station. Accordingly, robotic autonomy was the key mission-enabling factor. However, when a communication link was available, the Human Supervisor could provide high-level commands that optimized the robot's deployment strategy and the areas they undertook to explore. Accordingly, the robot autonomy was designed in a manner allowing for efficient, sparse, "supervision" by the human. In this "supervised autonomy" architecture, all the spatio-temporally dense path planning decisions were left for the robots to decide autonomously, while the operator could orchestrate the exploration strategy.

Acknowledgement

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00111820045. The presented content and ideas are solely those of the authors.

We want to thank all members of Team CERBERUS who contributed to the team's success. We extend our gratitude to all the SubT Community, and the DARPA team for the exciting challenge and collaborative community that was built.

References

1. D. Montgomery, "The Pentagon's $82 Million Super Bowl of Robots", *The Washington Post Magazine* (2021), `https://www.washingtonpost.com/magazine/2021/11/10/darpa-robot-competition/`.

2. Team CERBERUS (2022), `https://www.subt-cerberus.org/`.

3. M. Tranzatto, F. Mascarich, L. Bernreiter, C. Godinho, M. Camurri, S. Khattak, T. Dang, V. Reijgwart, J. Loeje, D. Wisth, others, Cerberus: Autonomous legged and aerial robotic exploration in the tunnel and urban circuits of the darpa subterranean challenge, *Field Robotics* 274–324 (2022).

4. P. De Petris, H. Nguyen, M. Dharmadhikari, M. Kulkarni, N. Khedekar, F. Mascarich, K. Alexis, Rmf-owl: A collision-tolerant flying robot for autonomous subterranean exploration, *arXiv preprint arXiv:2202.11055* (2022).

5. T. Dang, M. Tranzatto, S. Khattak, F. Mascarich, K. Alexis, M. Hutter, Graph-based subterranean exploration path planning using aerial and legged robots, *Journal of Field Robotics* 1363–1388 (2020).

6. M. Kulkarni, M. Dharmadhikari, M. Tranzatto, S. Zimmermann, V. Reijgwart, P. De Petris, H. Nguyen, N. Khedekar, C. Papachristos, L. Ott, R. Siegwart, M. Hutter, and K. Alexis, Autonomous teamed exploration of subterranean environments using legged and aerial robots, *IEEE International Conference on Robotics and Automation (ICRA)* (2022).

7. H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, J. Nieto, Voxblox: Incremental 3d euclidean signed distance fields for on-board mav planning, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2017).

8. L. Wellhausen, M. Hutter, Rough terrain navigation for legged robots using reachability planning and template learning, *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6914–6921 (IEEE, 2021).

9. T. Miki, L. Wellhausen, R. Grandia, F. Jenelten, T. Homberger, M. Hutter, Elevation mapping for locomotion and navigation using gpu, *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (submitted)* (2022).

10. T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, M. Hutter, Learning robust perceptive locomotion for quadrupedal robots in the wild, *Science Robotics* p. eabk2822 (2022).

11. M. Kamel, T. Stastny, K. Alexis, R. Siegwart. Model predictive control for trajectory tracking of unmanned aerial vehicles using robot operating system. *Robot operating system (ROS)* (Springer, 2017), 3–39.

12. S. Khattak, H. Nguyen, F. Mascarich, T. Dang, K. Alexis, Complementary multi–modal sensor fusion for resilient robot pose estimation in subterranean environments, *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1024–1029 (IEEE, 2020).

13. T. Schneider, M. Dymczyk, M. Fehr, K. Egger, S. Lynen, I. Gilitschenski, R. Siegwart, maplab: An open framework for research in visual-inertial mapping and localization, *IEEE Robotics and Automation Letters* 1418–1425 (2018).

14. J. Redmon, A. Farhadi, Yolov3: An incremental improvement, *arXiv preprint arXiv:1804.02767* (2018).

15. P. De Petris, S. Khattak, M. R. Dharmadhikari, G. G. Waibel, H. Nguyen, M. Montenegro, N. V. Khedekar, K. Alexis, M. Hutter, Marsupial walking-and-flying robotic deployment for collaborative exploration of unknown environments, *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (submitted)* (2022).