Nicholas Pey, Piotr Bonar, Xavier Sánchez, Carlota Criado & Judith Aguilera

Task #1 – Evaluation in Nascar Track

Q: Using this training ('N-S2-RL5-RD90'), how much can you increase the BaseDroneSpeed in the Nascar track while keeping a good performance (try only integer values)?

A: The highest integer value of BaseDroneSpeed that maintains good performance on the Nascar track using the N-S2-RL5-RD90 training is **3** (About 1 in every 3 starts, will it be able to successfully complete an entire lap).
Good performance is defined as the drone being able to consistently complete laps around the track without collisions.

Q: Train again the agent now with a Ray Length = 20. Is it better?

A: Yes, it is better. With a Ray Length of 20, the agent learns significantly faster — reaching a mean reward of **20.414 by step 45,000**. In contrast, with Ray Length 5, it only reaches a similar mean reward (**20.799**) at **step 100,000**. Regarding the highest integer value of BaseDroneSpeed that maintains good performance, the N-S2-RL20-RD90 training is also 3, which is the same as N-S2-RL5-RD90.

Q: Can you find a better configuration of the Ray Perception Sensor 3D modifying the ray length and the max ray's degrees? Store the ONNX that results from the best configuration and name it 'N-S2-BEST-RLx-RDy', where x is the length of the rays and y the max ray degrees used.

A: The best configuration is N-S2-BEST-RL15-RD45. It reaches the mean reward by step 40,000, however, similar to before, the max integer for BaseDroneSpeed is still 3.

Q: Write the parameters that give you better results and explain why you think this is the best combination.

A: First, our criteria for **better**:
- Reaches **mean reward ~20** in **fewer steps** (Step number in which >=20 is reached)
- Maintains good performance at **higher** BaseDroneSpeed (Highest possible BaseDroneSpeed)
- Bonus: Shows smooth movement and fewer collisions in Unity

Note: The table below shows the results for different speeds (last exercise)

Trial	Ray Length (x)	Max Ray Degree (y)	Best Speed (z)	Steps to converge	Speed at inference	Better? (Y/N)	Comment
1	5	90	-	NA	1	N	Only reaches a mean reward of ~11; insufficient for task
2	20	90	2	40,000	2	Y	Consistently reaches mean reward 20 after 40,000 steps
3	20	120	3	40,000	3	Y	Stable performance, occasional minor fluctuations
4	20	60	4	30,000	4	Y	Smoothest traversing and best obstacle avoidance. Highest consistency achieved
5	15	45	7	55	9	Y	
6	15	45	9	55000	9	Y	
7	20	90	5	55000	10	Y	
8	15	45	3	45	7	N	
9	15	60	5	35000	9	Y	
10	**20**	**60**	**5**	**40000**	**10**	**Y**	**THE BEST ONE**
11	30	90	-	NA	NA	N	Frequent collisions; consistently poor performance

After some thorough testing, which can be seen in the table above, we have decided that the best results are obtained with the following parameter combination (trained at):

- Speed = 5
- Ray Length = 20
- Max Ray's Degrees = 60

We have identified this combination of parameters as optimal for several key reasons. Primarily, it provides the most effective cost optimization, delivering efficiency without unnecessary complexity. Given the nature of the task, which involves navigating a relatively straightforward track, extensive maneuvering control is not essential. Instead, speed emerges as a critical factor for performance.

Hence, we selected a speed value of 5, as a lower setting would unnecessarily sacrifice valuable speed, without significantly improving the control required for this simple course. Additionally, our chosen values for "Ray Length" and "Max Ray Degrees" were carefully tuned to achieve faster convergence during training. These particular settings were the only ones that allowed us to reach a speed of 4, enabling better overall performance.

In summary, these parameters represent the optimal balance between speed and control, tailored specifically for tasks where agility is secondary to velocity.

Task #2 – Evaluation in Rally Track

> **Q:** Use the 'N-Sz-BEST-RLx-RDy' ONNX file you have generated in the Nascar track now in the Rally track (TrainingEnvRally). Adjust the BaseDroneSpeed at the same speed used for the training and **remember to adjust the Ray Perception Sensor** to the same parameters you used during the training. What do you observe? Now adjust the BaseDroneSpeed till you find a speed that allows the drone to complete all the laps.

A: When tested with the same parameters (speed 5, Ray Length 20, Max Ray Degree 60), the drone consistently crashed and couldn't complete any laps. This was expected, as the model was trained on a much simpler track. The Rally track has tighter curves and more complex navigation, which the model hadn't encountered.
After lowering the BaseDroneSpeed to 3, the drone was finally able to complete full laps, though with less stability. This suggests the model can adapt to some extent by slowing down, but it's clearly not optimized for the Rally environment.

> **Q:** Can you improve the behaviour by training in this track instead of using the ONNX trained in the Nascar track (you can modify the speed and Ray Perception Sensor parameters length and max rays degrees at your convenience for the training)? Name the resulting ONNX as 'R-Sz-BEST-RLx-RDy'. You can expect a maximum reward training in this track of ~19.5.

A: We followed the same training format as in Task #1, but this time on the Rally Track. To determine the best configuration, we focused on achieving the highest possible inference speed, as this typically indicates better generalization and less overfitting. At the same time, we valued fast convergence, since reaching optimal performance in fewer training steps reduces computational cost and reflects more stable learning dynamics.

The best result came from the following setup:

- **Ray Length**: 10
- **Max Ray Degree**: 60
- **Training Speed**: 3

Trial	Ray Length (x)	Max Ray Degree (y)	Training Speed (z)	Steps to converge	Speed on Inference	Better? (Y/N)
1	10	45	2	80	4	
2	10	45	3	145	5	
3	15	45	2	100	4	
4	20	45	2	35	4	
5	20	45	3	105	6	
6	10	60	2	55	3	
7	**10**	**60**	**3**	**25**	**13**	**BEST**
8	15	60	2	70	3	
9	20	60	2	NA	x	
10	20	60	3	40	4	
11	10	90	2	40	5	
12	10	90	3	50000	3	
13	20	90	2	40000	3	
14	20	90	3	80000	4	

This configuration converged in 25,000 steps, and during inference consistently reached a maximum speed of 13 (occasionally hitting 14, but not reliably).

> **Q:** Activate the Red drone in this track. Configure the White drone with the best Nascar ONNX and the Red drone with the best locally trained ONNX. What happens?

A: When both drones are deployed on the Rally track — the Red drone using the locally trained ONNX and the White drone using the Nascar-trained ONNX — the Red clearly shows better overall performance.

The Red drone runs at speed 3, so it's not the fastest, but it navigates the track consistently and without errors. As shown previously, it can even handle higher speeds without losing stability.

The White drone, on the other hand, is faster by default, but as explained earlier, it fails to complete the laps at its original speed due to poor control in this more complex environment. When we reduce its speed to 3, it does manage to finish laps, and occasionally it can even beat the Red drone. However, this success is not consistent, it varies from run to run, and its behavior remains unreliable.

In summary, while the White drone has higher top speed, the Red drone's robustness, adaptability, and potential to scale with increased speed make it the more reliable and effective option for the Rally track.

Task #3 – Evaluation on LeMans Track

> **Q:** Put together the new drone with the 'R-Sz-BEST-RLx-RDy' drone (assign it to the Red drone). Which one works better?

A: When testing both drones at the same speed (3)—which is the speed they were originally trained on—we observed that both were able to complete laps successfully. Surprisingly, the Red drone, which was trained on the simpler track, performed slightly better than the new one trained directly on the LeMans track.

Given these results, we decided to resume training on the Rally track and apply fine-tuning using the LeMans environment. It's important to note that the last row in the results table corresponds to this fine-tuned model, not a completely retrained one: it was initialized from the previously successful Rally-trained model.

Trial	Ray Length (x)	Max Ray Degree (y)	Training Speed (z)	Steps to converge	Speed on Inference	Better? (Y/N)
1	15	45	3	110	3	
2	20	45	3	95	3	
3	15	60	3	100	3	
4	20	60	3	115	4	
5	**10**	**60**	**3**	**150**	**4**	**BEST**

This fine-tuned version turned out to be the best-performing overall. It was trained with a speed of 3, a ray length of 10, and a maximum ray degree of 60, and thanks to its robustness, it was capable of increasing the inference speed to 4 while still completing laps reliably.

Interestingly, despite these improvements, the original Rally model (without fine-tuning) still outperformed the fine-tuned version in terms of stability and lap consistency at the same speed. While this might seem counterintuitive, it suggests that the base policy learned on the simpler track was already well-generalized, and that the fine-tuning phase may have introduced some level of overfitting or sensitivity to the LeMans layout. This insight highlights the need to carefully evaluate whether fine-tuning truly adds value in each specific context.

Note: Since fine tuning did not outperform previous results, you can find the second best model (.ONNX) in the .zip (L-S3-BEST-RL20-RD60)
Note: Linear Damping has to be set at 8