Parakeet: a C2 Framework for Efficient AI Capability Development

Valerie Lavigne and Mélanie Breton

October 1, 2020

Parakeet: a C2 Framework for Efficient AI Capability Development

Topic 7: Other C2 Related Research and Analysis

Concept Paper: Parakeet: a C2 Framework for Efficient AI Capability Development

Authors: Valérie Lavigne, Dr Mélanie Breton.

Abstract:

Command and Control (C2) concepts and practices have always co-evolved with communications and information technology. The recent advances in Artificial Intelligence (AI) have changed the way the information technology world performs. This requires C2 to quickly adapt in order to take full advantage of AI potential.

Deep Learning (DL) has been at the forefront of the recent advances in AI. As an example, in the field of computer vision, DL has enabled the development of object detection and classification models that can rival human capabilities in some domains [1]. However, DL algorithms are typically data hungry and they need to be trained on thousands of labelled samples. Annotated datasets that are relevant to military operations can be hard to come by. Although open source datasets are easily accessible, they rarely present the following characteristics that are prevalent in military operational context:

- Multiple sensor modalities (such as visible, infrared, thermal or hyper spectral);
- Degraded images (Noisy, blurred, low resolution);
- Labelled instances of specialized military equipment (not only everyday ordinary objects); and
- Aerial point of view (not only ground-based imagery).

To fill this gap, there is a need to complement datasets available from open source or allies, with data collection activities. Even though data collection itself requires significant efforts, the data labelling phase is often the most effort-intensive step of dataset creation. It is considered the main bottleneck holding back further AI adoption [2]. Labelling is challenging due to time constraint (it is a long and tedious process) and to subject matter expert availability for validating dataset quality.

In this paper, we propose the Parakeet framework which applies a C2 approach to labelled dataset building that creates the conditions necessary to AI capabilities development and operationalization success. This is done by ensuring an appropriate utilization of the available resources (operators, subject matter experts, time and computational power). We show that managing machine learning activities with a C2 framework can enable faster object detection and classification models development, which in turn will enable better C2 performance by providing timely situation analysis through the detection, recognition and tracking of objects and activities of interest.

References:

[1] Ranjan, Rajeev, et al. "Deep learning for understanding faces: Machines may be just as good, or better, than humans." IEEE Signal Processing Magazine 35.1 (2018): 66-83.

[2] Ilyas, Ihab and Lorica, Ben. "The quest for high-quality data." June 18, 2019. O'Reilly Radar / AI & ML. https://www.oreilly.com/radar/the-quest-for-high-quality-data/, accessed April 2, 2020.

Parakeet: a C2 Framework for Efficient AI Capability Development

Valérie Lavigne, Mélanie Breton

Introduction

Deep Learning (DL) has been at the forefront of the recent advances in AI. In the field of computer vision, DL has enabled the development of object detection and classification models that can rival human capabilities in some domains [1]. However, DL algorithms are typically data hungry; they need to be trained on thousands of labelled samples. Annotated datasets that are relevant to military operations can be hard to come by.

To fill this gap, there is a need to complement datasets available from open source or allies, with data collection activities. Even though data collection itself requires significant efforts, the data labelling phase is often the most effort-intensive step of dataset creation. Nowadays, this step is considered as the main bottleneck holding back further AI adoption [2]. Labelling is challenging due to time constraint (it is a long and tedious process) and to subject matter expert availability for validating dataset quality.

In this concept paper, we argue that AI capability development process needs to be organized in a deliberate way to ensure efficiency (optimize human effort on dataset labelling), information quality assurance (build trust, ensure model validity and dataset integrity), and performance (measure model performance within the intended operating conditions).

As a first step to achieve this goal, we propose the Parakeet framework which applies a C2 approach to labelled dataset building that creates the conditions necessary to AI capabilities development and operationalization success. This is done by ensuring an appropriate utilization of the available resources (operators, subject matter experts, time and computational power).

Finally, we end the paper with our thoughts on how C2 will benefit from the new capabilities afforded by leveraging the AI potential.

Developing AI Capabilities

AI Development Strategy

As part of the AI development cycle, it is important to clearly define the capabilities that need to be developed. This can be done in an initiation phase where the purpose of the model should be defined, along with quantitative performance goals, and the intended operating conditions. These can vary widely depending on the task to perform. Counting vehicles in a parking lot will require a lower level of precision on object localization than tracking individual vehicle movement patterns, even though both are related to vehicle detection in sensor feeds. The operating conditions must be considered because deep learning algorithms are sensitive to background context and a model developed using a dataset collected under summer conditions may not perform as intended when we apply it in a winter context where the presence of snow affects lighting conditions and image contrast.

Validating the end results will be part of the AI development cycle as well. Although this step usually comes at the end, it needs to be considered at the beginning since gathering validation data is likely to involve the same efforts as what is needed for training data, and it will be more efficient to do both at the same time. The metrics for evaluating the end results should also be defined early in the process as they will be used throughout the model training phase.

Between the initiation and validation phases we need to build detection models and to curate the datasets required to train them. These two steps usually form the most effort-intensive part of the process, and our proposed Parakeet framework aims at efficiently managing those resources.

Curating Datasets

Labelled datasets that are relevant to military operations are often hard to get by. Although open source datasets are easily accessible, they rarely present the following characteristics that are prevalent in military operational context:

- Multiple sensor modalities (such as visible, infrared, thermal, or hyperspectral);
- Degraded images (noisy, blurred, low resolution);
- Labelled instances of specialized military equipment (not only everyday ordinary objects); and
- Multiple platforms point of view (air-based and ground-based imagery).

At the heart of the AI development cycle presented above is the need for proper training and evaluation data to develop a model and assess its performance. Three approaches contribute to this goal:

- **Exploitation:** Leverage external sources, align and integrate annotations;
- **Collection:** Capture from sensors and annotate data; and
- **Augmentation:** Apply randomized transformations, generate synthetic data.

Exploitation of existing dataset should always be considered first. However, we should not underestimate the level of effort required to repurpose and integrate those existing datasets. First, we must identify and gain access to those datasets. Then, in order to integrate them, we would need to spend time to align semantically the annotations and the ontology for the new purpose.

Data augmentation strategies will make an existing dataset better, however, they do not replace real imagery.

Regarding data generation, the research community has not demonstrated so far that models training on synthetic datasets only generate well enough to real situations. In any case, a real-world dataset will still be required for validating the model’s results in a realistic context. Finally, after evaluating what is available, i.e. assessing the dataset coverage and biases, it should be complemented by performing new collection activities. Both the collection and exploitation approaches are likely to require significant labelling efforts to ensure the proper labels are assigned to the data elements.

Finally, after evaluating what data is available, i.e. assessing the dataset coverage and biases, it should be complemented by performing new collection activities. Both the collection and the exploitation approaches are likely to require significant labelling efforts to ensure the proper labels are assigned to the data elements.

C2 Applied to Labelling Missions

Labelling Cycle

Taking a C2 perspective, we can consider the dataset annotation cycle like a targeting mission cycle. Indeed, in order to better convey the steps involved in the process of dataset curating and model training, we can align the labelling cycle steps to those of the Joint Targeting Cycle [3], as shown in Figure 1.

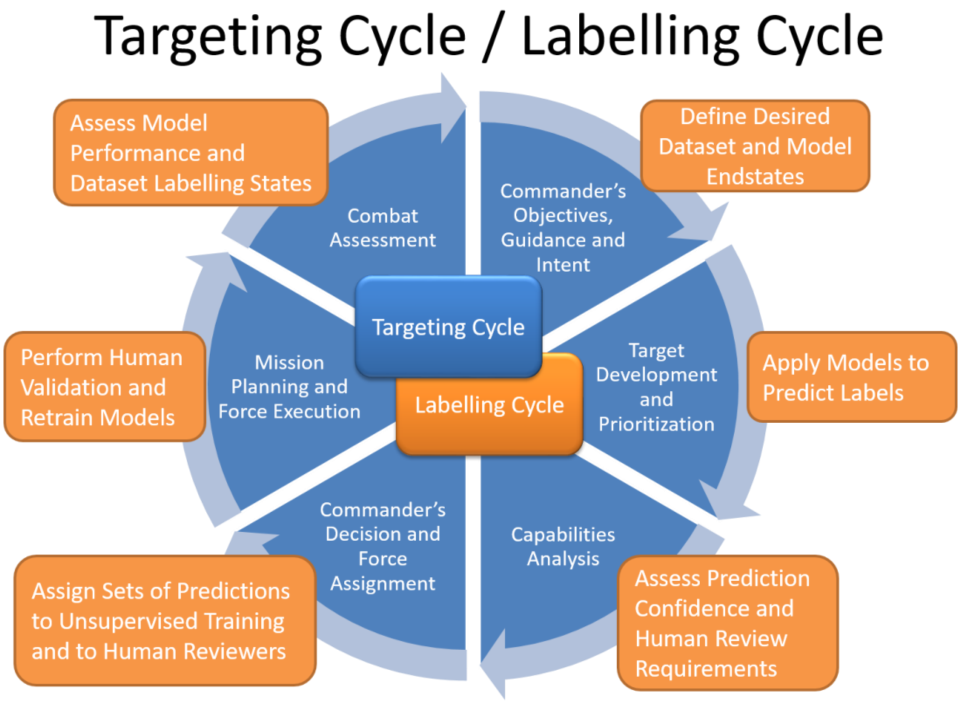

Figure 1: Parakeet's Labelling Cycle Mapped to the Targeting Cycle.
(Targeting Cycle Source: B-GJ-005-309-FP-001, CFJP 3.9 Targeting [3])

This cycle involves the following activities that the reader who has knowledge of the targeting process will easily recognize.

Commander's Objectives, Guidance and Intent: At this point, we define the desired end goals of our labelling mission in terms of dataset labelling coverage and model performance requirements. These are dependent on the purpose of the labelling tasks. We determine minimum quantities that must be present for each type of object of interest and select expected label names. As these labels are mission-related, they will vary across different tasks, depending on the level of details required, and the selected trade-off between detection accuracy and development effort. Finally, we set minimal acceptable confidence level to allow for unsupervised learning and configure the training parameters of the model.

Target Development and Prioritization: We identify the labelling priorities by making an inventory of the labelling information we already have in order to highlight what is missing and should be tackled first. If available, we apply existing models to compute predictions for unlabelled data. We assess prediction confidence levels and the human review requirements for all data elements, in order to produce a prioritized list.

Capabilities Analysis: We evaluate models' capability to label objects in an unsupervised fashion. We determine the availability and workload for human reviewers, indicating the number of data elements that can be validated in this iteration.

Commander's Decision and Force Assignment: We allocate limited human validation resources, i.e. assign sets of predictions either to humans for review, or to the unsupervised training set for the next model retraining phase. We decide which data elements will be set aside for the validation set. We distribute labelling set orders to human reviewers.

Mission Planning and Force Execution: At this step, subject matter experts perform the actual validation and models are retrained as needed.

Combat Assessment: We compute metrics to assess model performance and dataset labelling states. If the end goals are not met, the cycle is repeated. This is also a step where we can evaluate if we need to refocus labelling on certain types of objects for which we have few examples, akin to setting and meeting Priority Information Requests for the Intelligence cycle.

Our Collection and Data Labelling Experiences

Through our R&D work on deep learning capabilities, we have performed collection, integration, and labelling of various datasets. In October 2019, DRDC led a data collection activity in collaboration with the 5e Régiment d'artillerie légère du Canada (5e RALC) to build a video dataset featuring multiple Canadian armoured vehicles. More than 16 sensors were used to collect imagery, including intensified, visible and infrared cameras, located either on the ground or in the air (drones). This led to more than 2.5 Tera bytes of data acquired in a single day. Labelling this dataset represented a significant effort. A second data collection activity in February 2020, also done in collaboration with the 5e RALC and the 1er Batallion Royal 22e Regiment led to more than 4 Terabytes of data acquired by more than 25 sensors on site. Obviously, a strategy for labelling these datasets needed to be developed.

In a related project, DRDC coordinated a labelling session over a month in November 2018 on an aerial video dataset. Two full time military personnel from the CFB Valcartier Personnel Awaiting Training (PAT) Platoon identified around 121 000 objects in more than 40 000 images over the period of 4 weeks. In comparison, the open source ImageNet dataset [4] has more than 1.2 million labelled images. Working at the PAT Platoon labelling rate, the required labelling effort to equal a training set the same size as the ImageNet dataset would take more than 30 months.

While developing a prototype for car colour labelling, we experimented with a different approach where we leveraged an iterative model training process. First, we spent 43 hours to manually label the first 8 000 images. Then we applied a semi-automated approach on the next 7 000 images which were pre-labelled by the model. After that, the human reviewing process took less than an hour. This represented a significant efficiency improvement (more than 40 times faster labelling), which inspired us to develop the Parakeet concept. The same exercise was done a second time while building a military vehicle dataset. This time, we first labelled 8000 images in 8 hours. Then, the reviewing process of more than 46 000 images took less than 2 hours.

Parakeet Concept

This section describes the Parakeet concept and its main components, which leverage and automate an active learning strategy for dataset labelling and model training. More specifically, Parakeet improves on traditional labelling approaches in three ways:

- **Efficient validation**: Parakeet predicts object bounding box localization. A human labeller simply validates the boxes, rather than drawing them from scratch.
- **Unsupervised learning**: Parakeet automatically includes automatically objects with high prediction confidence levels in the training set.
- **Active learning**: Parakeet selects objects bounding boxes with low predictions confidence level for human validation.

Figure 2 shows the full Parakeet concept for multimodal image and video datasets. Working with video datasets adds the possibility to leverage frame interpolations. Multimodal datasets can include a mix of image/video modalities (such as visible and infrared). In this case, we can leverage a detection model trained on a modality (visible) to label data from another modality (infrared). Moreover, if data was captured synchronously by two different sensors, we can leverage this information to transfer the labelling information between sensors modalities.

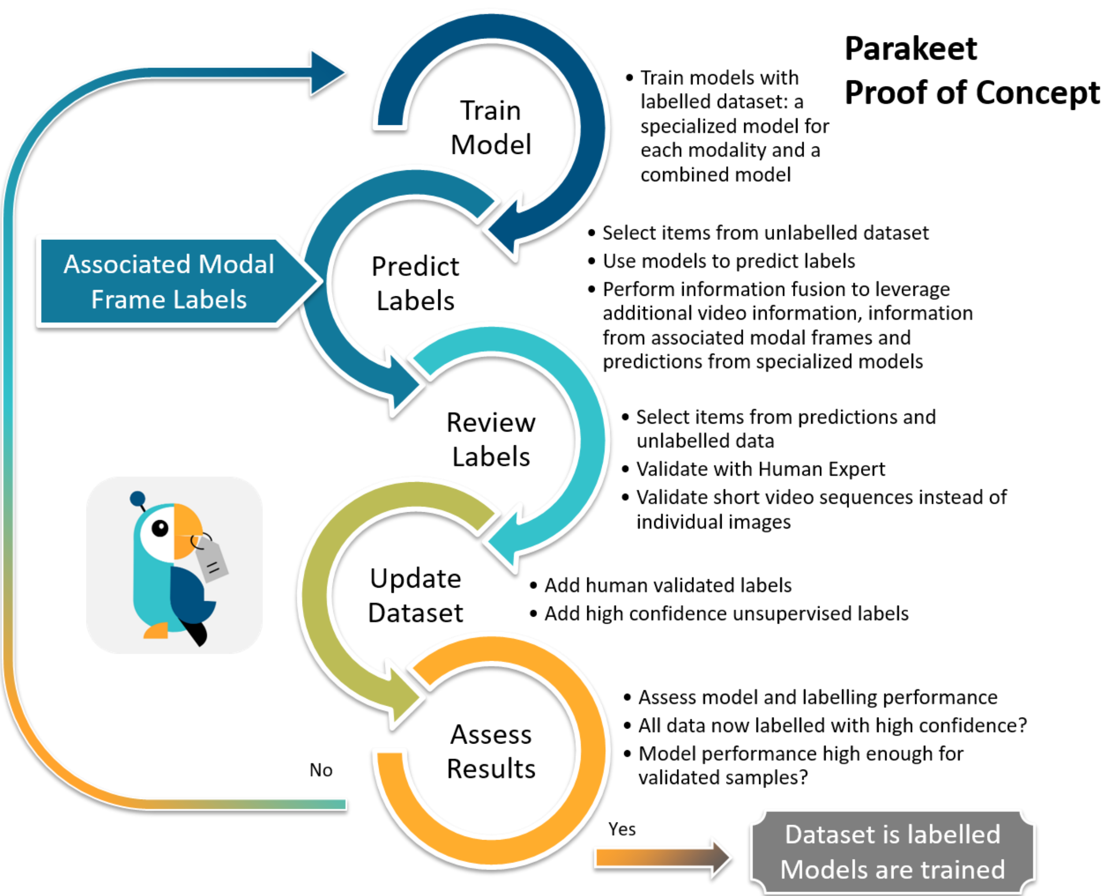

Figure 2: Full Parakeet Prototype Concept.

In Figure 3, we can see the seven main components which compose the Parakeet framework:

1. **Annotation Repository**: central repository for storing and managing dataset information, including the annotations information;
2. **Detection and Classification Models**: an ensemble of models that perform object detection and classification predictions;
3. **Model Training Module**: a set of algorithms that trains the detection and classification models using the available training datasets;
4. **Prediction Module**: performs object detection and classification of all the available data using the models trained, but also leverages other predictions strategies such as associated modality transfer, visual object tracking and frame interpolation;
5. **Human Validation Module**: a graphical user interface that allows a human operator to create new annotations, as well as to validate and correct the proposed object labels and bounding boxes;
6. **Performance Assessment Module**: measures both the models' performance and the dataset labelling progress, in order to inform decision-making on whether the dataset labelling and models training processes are satisfactory; and
7. **Data Triage Module**: this is the component responsible for handling all the C2 coordination aspects of the labelling process. This module receives information from the other modules, coordinates them and decides where and how data should be allocated to each subset.

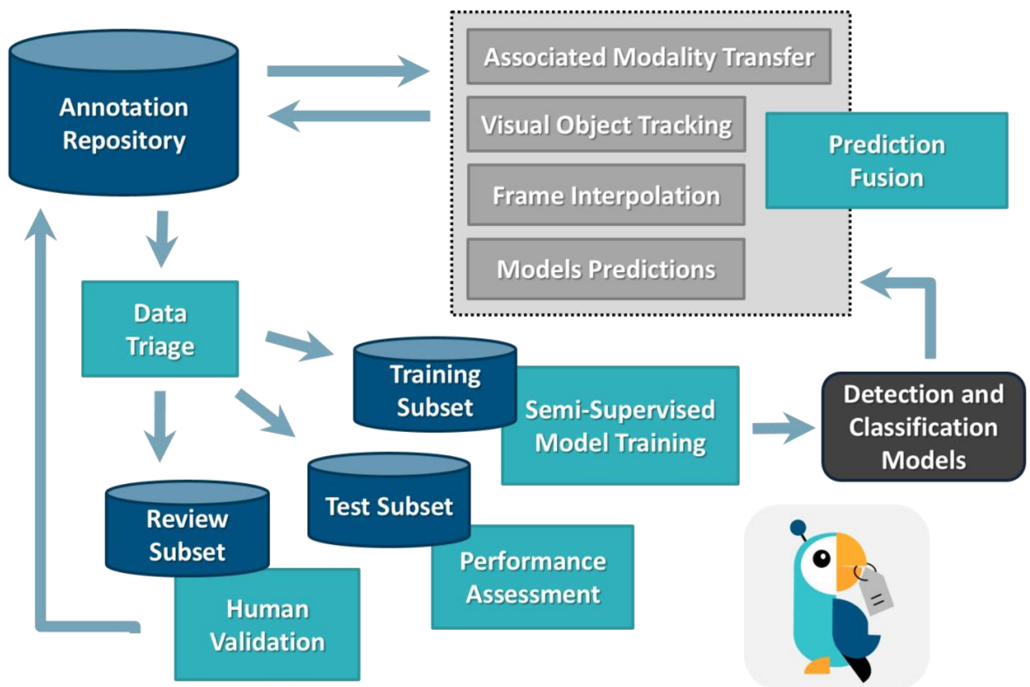

Figure 3: Parakeet framework components.

AI Impact on C2

Command and Control concepts and practices have always co-evolved with communications and information technology. The recent advances in Artificial Intelligence (AI) have changed the way the information technology world performs. This requires C2 to quickly adapt in order to take full advantage of AI potential.

Applying C2 to shape our AI approach at the enterprise level will determine the level of force multiplier we can reach and the extent to which these capabilities are possible over time. In return, new AI capabilities will influence the conduct of C2. We expect that AI can be leveraged in multiple ways. Better object detection from sensor feeds can inform a more precise and complete common operating picture, enabling timely situation analysis that contributes to battle space management and awareness. Recognizing situation changes and following situation evolution is critical to ensure C2 agility to react. AI has the potential to accelerate the decision-action cycle by detecting those situations automatically. Moreover, there are many small areas of improvements where AI can add efficiency and partial automation resulting in faster response times requiring fewer human resources or allowing personnel to spend their time on more important tasks.

To reach that point, we need to address key information assurance about training datasets to ensure we can identify inconsistent, outdated, incorrect information, as well as misinformation. This is the basis for establishing trust in the AI capabilities that are created and it requires a strategic approach to managing AI development.

Of course, AI will not provide all the answers, but we can expect that those who learn how to best leverage AI technologies will develop a competitive advantage over their adversaries.

Conclusion

In this paper we presented the Parakeet framework, which exploits active learning, unsupervised learning, and efficient prediction validation to enable faster dataset labelling and model training.

We showed that managing machine learning activities with a C2 framework can enable faster object detection and classification model development, which in turn will enable better C2 performance by providing timely situation analysis through the detection, recognition and tracking of objects and activities of interest.

Ensuring quality sensor detection feeds is only a first step. We expect that C2 will both shape and employ AI. There is a need to reflect on how we can best exploit the new AI capabilities to support better situation awareness and decision-making. We need an AI development strategy that is flexible and adaptable to new problems.

References

[1] Ranjan, Rajeev, et al. "Deep learning for understanding faces: Machines may be just as good, or better, than humans." IEEE Signal Processing Magazine 35.1 (2018): 66-83.

[2] Ilyas, Ihab and Lorica, Ben. "The quest for high-quality data." June 18, 2019. O'Reilly Radar / AI & ML. https://www.oreilly.com/radar/the-quest-for-high-quality-data/, accessed April 2, 2020.

[3] Canada. Department of National Defence. B-GJ-005-309-FP-001. Canadian Forces Joint Publication 3-9 Targeting. Ottawa: DND Canada, 2008.

[4] ImageNet. ImageNet Summary and Statistics. http://image-net.org/about-stats. accessed 5 June 2020.