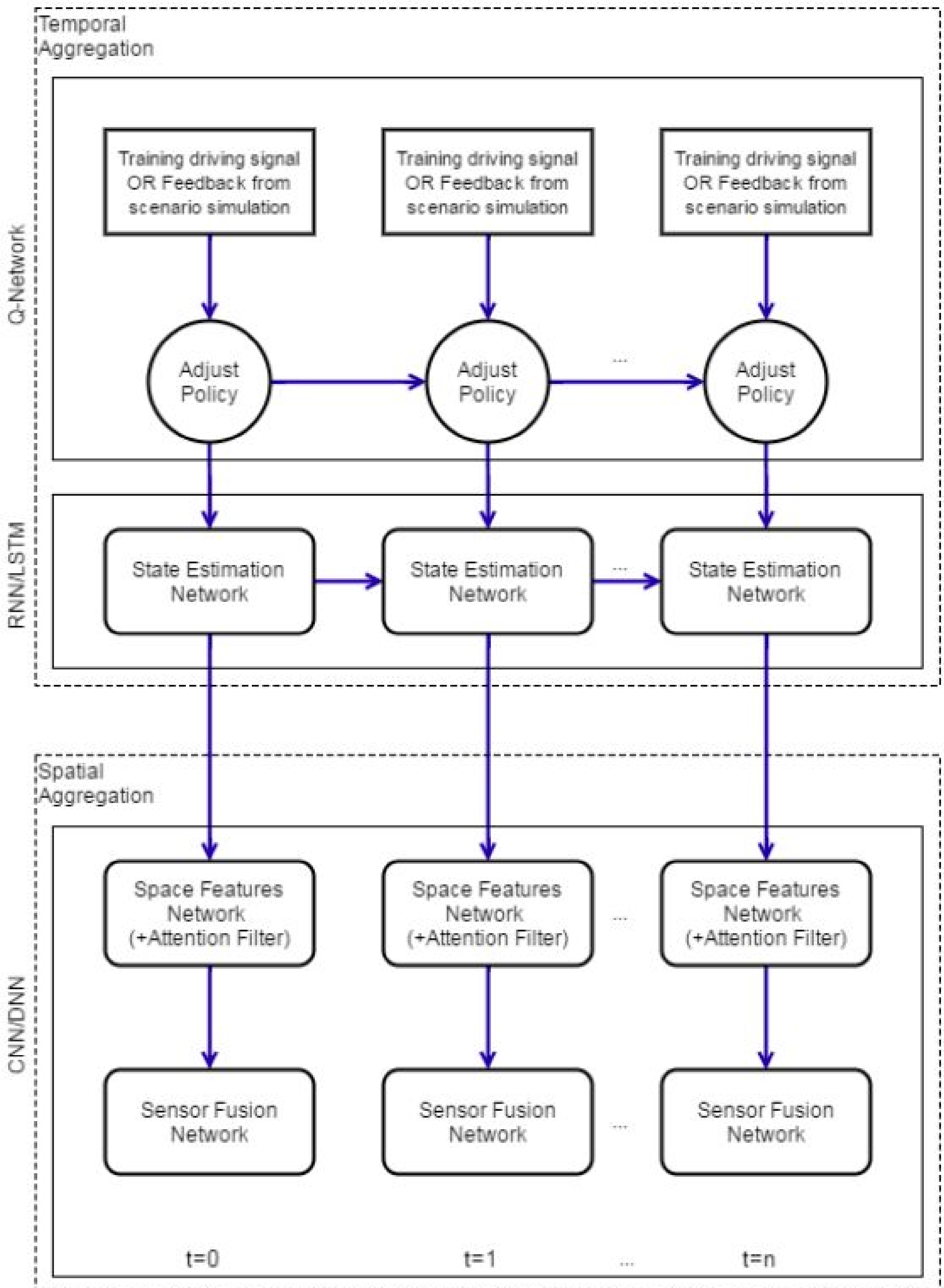
Deep Reinforcement Learning for Autonomous Driving

In recent years, the field of Computer Science has been advancing rapidly particularly with respect to Artificial Intelligence - which while having been a field of research for many years, is now allowing developers come up with real world solutions. A prime example of this radical jump in technological progress is the field of autonomous driving. Seemingly straight out of science fiction, the idea of autonomous driving is one that has enticed leaders in technology for many years, and recent developments in the sub-field of Machine Learning - have made this a reality: Tesla's autonomous cars being a prime example of this. Nonetheless, for a problem as complex as driving - which proves a challenge even for humans (as seen by the restrictions on under-age driving), there lies huge scope for further improvement in the realm of autonomous driving. Hence Ahmad El Sallab, Mohammed Abdou, Etienne Perot and Senthil Yogamani, researches and developers at Valeo Vision systems chose to improve existing learning algorithms for driving by proposing a framework relying principally on 'Deep Reinforcement Learning'. This choice was motivated primarily by the success that such a style of learning had had for autonomous agents playing Atari games that are similar in computational complexity. Moreover, autonomous driving provided the ideal conditions for application of 'Reinforcement Learning' as due to the strong dependence of this task on interactions with the environment. The researchers further justified their motivations, by stressing that this interaction also makes it difficult to model the problem as Supervised Learning problem.

The first section of the paper dealt with the higher-level overview of what tasks would be involved in the creation of an autonomous driving agent. The authors chose to illustrate how the plethora of tasks involved in this monolithic process could be broken down into three main sections: Recognition, Prediction and Planning; to set up the groundwork for the various algorithmic and logical choices made in the final proposal for a new framework for autonomous driving. The first set of tasks that could broadly be grouped as 'Recognition' dealt with "identifying components of the surrounding environment" (1) (equivalent to tasks that could be accomplished by a human child of 2-3 years). This tasks involved a simple recognition of static objects, thus being the creation of a one-to-one mapping from object to 'what it is', devoid of any integration of past information. An example of such a task would be recognizing a bus or a traffic sign while driving. The researchers mentioned how this was the easiest of the three groups of tasks, not because of the trivial nature of the task - because the task in itself is not trivial in any manner; but rather due to the existence of many algorithms that can accomplish this to near-human perfection since the development of AlexNet - a Convolutional Neural Network that at the annual ImageNet challenge proved the efficacy of this style of networks for the problem of object recognition in images. The second major set of tasks mentioned in this paper was Prediction, in particular "predicting future states of the environment" (4 - 16 year old) (1). This task is significantly more complex as it involved tracking dynamic objects and integration information over time as humans themselves use the past movements of an object to ascertain its future movements. This problem is made particularly challenging by the lack of perfect information - as objects that are being tracked may disappear from vision or be out of range for the LIDAR (or be in some way impossible to track for certain sensors). This lack of perfect information motivated the researchers to use Recurrent Neural Networks for this section (specifically, Long-Short Term Memory (LSTM) networks) as LTSM networks function much like human short term memory - thus being well-suited to classifying, processing and prediction actions despite time lags of unknown size between important events. The final and most complex stage in the process of autonomous driving is 'Planning' - the objective being to "plan the future sequence of driving actions that will enable the vehicle to navigate successfully (16+ years old - ability increases with time)" (1). The difficulty here lies in the complex task of planning actions in a dynamic environment that requires integration over time (both past and future) to choose the best policy (set of actions) for the autonomous driving agent. It is this task that the framework accomplishes using 'Deep Reinforcement Learning'.

The paper then went into depth regarding the motivation behind the various algorithmic decisions made for the framework, which have been briefly surveyed here. The use of 'Reinforcement Learning' was driven by the nature of the problem i.e. a multi-agent problem with interaction that can be quantified easily into rewards and penalties. This in conjunction with the demonstration of human-level control on Atari Games encapsulates the thought processes behind choosing this form of learning. In addition, the framework utilizes a 'Deep Attention Reinforcement Learning' Algorithm which strives to mimic the way humans process information while driving (focus on key objects like other cars and dividers, while paying less or no attention to rocks or the clouds), by filtering out irrelevant data or weighting it accordingly. This filtering helps reduce the dimensionality of the data, reducing the computational load of the algorithm - thus making it more efficient. Framing this as a learning process that occurs in parallel to the actual driving, allows the attention filters to improve while not adding load on the processing required for the actual autonomous driving.

Figure 1.0: Visualization of the Proposed Framework



The final section of the paper finally details the framework that the researchers proposed, but it is easy to see why this would have merit - considering the justification provided for algorithmic choice of Deep Reinforcement Learning, Attention Filters and Recurrent Neural Networks. The first stage is the spatial aggregation network which contains of two networks: Sensor Fusion and Spatial Feature Recognition (self-explanatory). Sensor Fusion refers to the collection of environment data, includes information not only about external environment state (objects, their position, velocity etc.) but also information about the car state i.e. its position, orientation, velocity and acceleration. The critical difficulty here is fusing the data in a robust manner. This fused sensor data is then piped to the spatial feature recognition network. This begins by first extracting basic low-level features of the environment, which are piped to a attention filter that reduces the dimensionality of the information so that the final 'Convoluted Neural Network' that extracts deeper representations (e.g. signpost, bus etc.) has lesser computational load on it. The second major stage of the framework is the temporal aggregation that on parsing the information from the Spatial Aggregation Network, uses Deep Reinforcement Learning coupled with Recurrent Neural Networks to effectively predict the future states of the surrounding environment and come up with the best policy $\pi(s)$.

To conclude, this framework was tested in TORCS (The Open-source Racing Car Simulator) with the Simulated Car Racing add (which gives access to the car controls, like steering, velocity, acceleration, brakes, and car states, like position, velocity, acceleration, fuel level, etc.) yielded favourable results, thus showing merit for this Reinforcement Learning based framework. The research team mentioned that future work would involve deploying the algorithm for a car in a simulated environment (where sensors and actuators are artificially controlled) to allow the framework to be tested more robustly.

Bibliography

1. Sallab, Ahmad El, et al. "Deep Reinforcement Learning Framework for Autonomous Driving." ArXiv, 8 Apr. 2017, arxiv.org/pdf/1704.02532.pdf.