Bagging Deep Q-Networks

Jamis Johnson, Lance Legal, Angus Ding
Department of Computer Science
Columbia University
New York, NY
[jamis.johnson,lwl2110,ad3180]@columbia.edu

Abstract

Recent advancements by DeepMind in leveraging convolutional neural networks to represent the state space in Q-learning have dramatically improved the performance of algorithmic game playing, outperforming nearly all competing algorithms across 49 Atari 2600 games. In this paper, we propose a straight forward, distributed method for increasing the efficacy of deep Q-network agents thorugh the well-known ensemble algorithm, boostrap aggregation. We have developed software to orchestrate the deployment, training, and testing of this method on a cluster of GPU optimized, Amazon Web Service EC-2 virtual machines. We provide results demonstarting the effectiveness of training time on performance as well as ensemble size to performance.

1 Overview

Q-networks have been shown to outperform 43 state-of-the-art agents across a diverse array of 49 Atari 2600 games, often by extreme margins.

1.1 Deep Q-networks

Q-learning is a long-standing, model-free reinforcement learning algorithm used to find an optimal action-selection policy[1]. Deep Q-networks are a novel approach to Q-learning in which deep convolutional neural networks are used to reduce high-dimensional raw input to a set of possible actions[2]. The convolutional neural network approximates the optimal action-value function

$$Q^*(s,a) = \max_{\pi} \mathbb{E}\left[r_t + \gamma r_{t+1}\gamma^2 r_{t+2} + \ldots | s_t = s, a_t = a, \pi\right]$$

which is the maximum sum of discounted rewards at time t. It is quite common to approximate Q^* using a linear approximator, but in this case a nonlinear convolutional neural network is used such that, given θ weights, $Q(s,a;\theta) \approx Q^*(s,a)$. Spatial, convolutional neural networks have drastically increased the accuracy of image recognition tasks in recent years are therefore a natural fit as a state-space function approximator of raw image data[3,4,5,6,7].

Memory replay

DeepMind emplyed a biologically-inspired memory replay mechanism that was integral to the success of Q-networks. Experiences, $e_t = (s_t, a_t, r_t, s_{t+1})$, are stored at every time step t, where data set $D_t = \{e_1, \ldots, e_t\}$. Q-learning updates are applied during learning on uniformly sampled experiences, with the loss function

$$L_i(\theta_i) = \mathbb{E}_{(s,a,r,s')\sim U(D)} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta) \right)^2 \right]$$

θ_i are the parameters to the Q-network and θ_i^- are the parameters used in the previous iteration to calculate the target.

Network architecture

TODO

1.2 Bagging

Bootstrap aggregation is a standard technique for improving the stability and performance of statistical classifiers[8]. It is commonplace to employ this method when modeling using decision trees, but has also been shown to produce good results when applied to artificial neural networks[9,10].

Bagging is a process of averaging across an ensemble of classifiers $f^{*n}(x)$, where $n \in [1, N]$ and training data consists of N Bootstrap samples, such that

$$f_{bag}(x) = \frac{1}{N} \sum_{n=1}^{N} f^{*n}(x)$$

An intuitive way to think about bagging is: "the wisdom of crowds."

Bagging Q-networks

In a supervised learning environment you would train N neural networks, f_{*n}, on bootstrapped data sets, $Z_n = \{(x_1, y_1), \ldots (x_b, y_b)\}_n$. In the case of Q-networks we do not have

1.3 Alternate Approaches Considererd

Let's not opine too long on these alternate approaches as they are second to bagging

Multi-GPU Training Performance Speedup

Krhizevsky(spelling)'s "neat trick" to optimizing the traning of convolutional neural networks across multiple GPUs[TODO] was initially appealing but fell by the wayside when it became clear that saturating the memory and compute power of a single NVIDIA K520 GPU would not be straight forward. We still feel there is room for work to be done here in optimizing DeepMind's now dated deep Q-network Torch code to automatically utilize multiple, single-machine GPUs.

Novel Exploratory Parameter

2 Implementation

2.1 Cluster Architecture

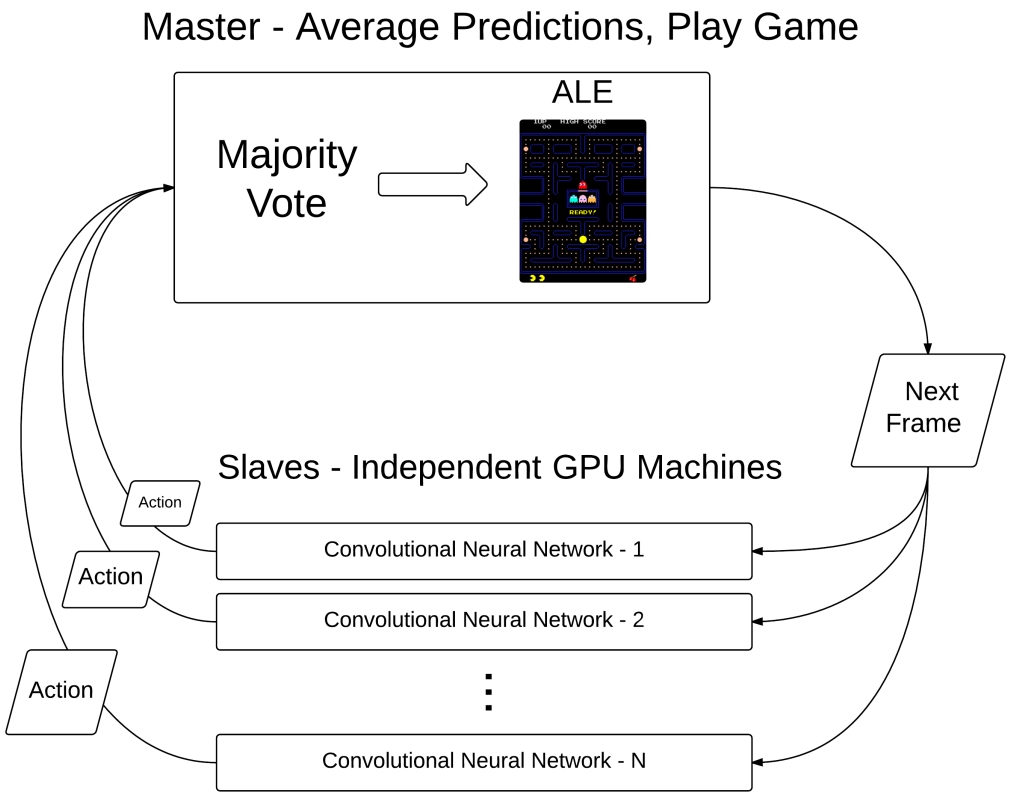

Figure 1: Cluster Architecture

Table 1: TODO: Title

Config	Space	Breakout	Pacman
3x6	6	5	5
3x10	6	5	5
3x18	6	5	5

3 Results

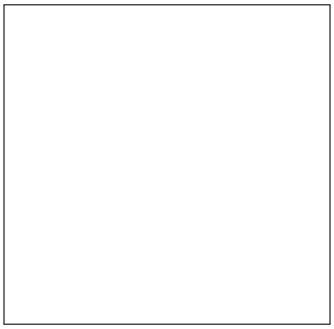

Figure 2: Effect of Ensemble Size on Performance.

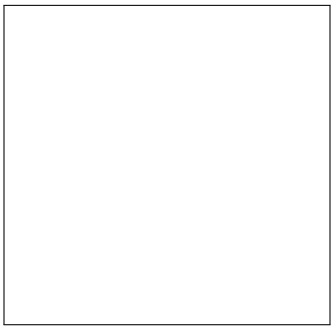

Figure 3: Effect of Training Time on Ensembles.

4 Conclusion

Acknowledgments

Thanks Mom

References

[1] Watkins, Christopher JCH, and Peter Dayan. "Q-learning." *Machine learning* 8.3-4 (1992): 279-292.

[2] Mnih, Volodymyr, et al. "Human-level control through deep reinforcement learning." *Nature* 518.7540 (2015): 529-533.

[3] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.

[4] Szegedy, Christian, et al. "Going deeper with convolutions." *arXiv preprint arXiv:1409.4842* (2014).

[5] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[6] He, Kaiming, et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification." *arXiv preprint arXiv:1502.01852* (2015).

[7] Sermanet, Pierre, et al. "Overfeat: Integrated recognition, localization and detection using convolutional networks." *arXiv preprint arXiv:1312.6229* (2013).

[8] Breiman, Leo. "Bagging predictors." *Machine learning* 24.2 (1996): 123-140.

[9] Schwenk, Holger, and Yoshua Bengio. "Boosting neural networks." *Neural Computation* 12.8 (2000): 1869-1887.

[10] Ha, Kyoungnam, Sungzoon Cho, and Douglas MacLachlan. "Response models based on bagging neural networks." *Journal of Interactive Marketing* 19.1 (2005): 17-30.

[] Krizhevsky, Alex. "One weird trick for parallelizing convolutional neural networks." *arXiv preprint arXiv*:1404.5997 (2014).

[] Dean, Jeffrey, et al. "Large scale distributed deep networks." *Advances in Neural Information Processing Systems*. 2012.

[] Collobert, Ronan, Koray Kavukcuoglu, and Clment Farabet. "Torch7: A matlab-like environment for machine learning." *BigLearn, NIPS Workshop*. No. EPFL-CONF-192376. 2011.

5 Submission of papers to NIPS 2015

NIPS requires electronic submissions. The electronic submission site is

`http://papers.nips.cc`

Please read carefully the instructions below, and follow them faithfully.

5.1 Style

Papers to be submitted to NIPS 2015 must be prepared according to the instructions presented here. Papers may be only up to eight pages long, including figures. Since 2009 an additional ninth page *containing only cited references* is allowed. Papers that exceed nine pages will not be reviewed, or in any other way considered for presentation at the conference.

Please note that this year we have introduced automatic line number generation into the style file (for LaTeX 2_ε and Word versions). This is to help reviewers refer to specific lines of the paper when they make their comments. Please do NOT refer to these line numbers in your paper as they will be removed from the style file for the final version of accepted papers.

The margins in 2015 are the same as since 2007, which allow for $\approx 15\%$ more words in the paper compared to earlier years. We are also again using double-blind reviewing. Both of these require the use of new style files.

Authors are required to use the NIPS LaTeX style files obtainable at the NIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

5.2 Retrieval of style files

The style files for NIPS and other conference information are available on the World Wide Web at

`http://www.nips.cc/`

The file `nips2015.pdf` contains these instructions and illustrates the various formatting requirements your NIPS paper must satisfy. LaTeX users can choose between two style files: `nips15submit_09.sty` (to be used with LaTeX version 2.09) and `nips15submit_e.sty` (to be used with LaTeX2e). The file `nips2015.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own. The file `nips2015.rtf` is provided as a shell for MS Word users.

The formatting instructions contained in these style files are summarized in sections **??**, **??**, and **??** below.

6 General formatting instructions

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing of 11 points. Times New Roman is the preferred typeface throughout. Paragraphs are separated by 1/2 line space, with no indentation.

Paper title is 17 point, initial caps/lower case, bold, centered between 2 horizontal rules. Top rule is 4 points thick and bottom rule is 1 point thick. Allow 1/4 inch space above and below title to rules. All pages should start at 1 inch (6 picas) from the top of the page.

For the final version, authors' names are set in boldface, and each name is centered above the corresponding address. The lead author's name is to be listed first (left-most), and the co-authors' names (if different address) are set to follow. If there is only one co-author, list both author and co-author side by side.

Please pay special attention to the instructions in section **??** regarding figures, tables, acknowledgments, and references.

7 Headings: first level

First level headings are lower case (except for first word and proper nouns), flush left, bold and in point size 12. One line space before the first level heading and 1/2 line space after the first level heading.

7.1 Headings: second level

Second level headings are lower case (except for first word and proper nouns), flush left, bold and in point size 10. One line space before the second level heading and 1/2 line space after the second level heading.

7.1.1 Headings: third level

Third level headings are lower case (except for first word and proper nouns), flush left, bold and in point size 10. One line space before the third level heading and 1/2 line space after the third level heading.

8 Citations, figures, tables, references

These instructions apply to everyone, regardless of the formatter being used.

8.1 Citations within the text

Citations within the text should be numbered consecutively. The corresponding number is to appear enclosed in square brackets, such as [1] or [2]-[5]. The corresponding references are to be listed in the same order at the end of the paper, in the **References** section. (Note: the standard BIBTEX style `unsrt` produces this.) As to the format of the references themselves, any style is acceptable as long as it is used consistently.

As submission is double blind, refer to your own published work in the third person. That is, use "In the previous work of Jones et al. [4]", not "In our previous work [4]". If you cite your other papers that are not widely available (e.g. a journal paper under review), use anonymous author names in the citation, e.g. an author of the form "A. Anonymous".

8.2 Footnotes

Indicate footnotes with a number[1] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).[2]

8.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction; art work should not be hand-drawn. The figure number and caption always appear after the figure. Place one line space before the figure caption, and one line space after the figure. The figure caption is lower case (except for first word and proper nouns); figures are numbered consecutively.

Make sure the figure caption does not get separated from the figure. Leave sufficient space to avoid splitting the figure and figure caption.

You may use color figures. However, it is best for the figure captions and the paper body to make sense if the paper is printed either in black/white or in color.

8.4 Tables

All tables must be centered, neat, clean and legible. Do not use hand-drawn tables. The table number and title always appear before the table. See Table **??**.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

[1] Sample of the first footnote

[2] Sample of the second footnote

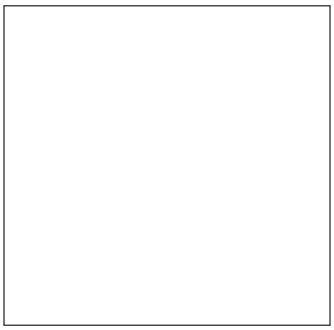

Figure 4: Sample figure caption.

Table 2: Sample table title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

9 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

10 Preparing PostScript or PDF files

Please prepare PostScript or PDF files with paper size "US Letter", and not, for example, "A4". The -t letter option on dvips will produce US Letter files.

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NIPS. Please see `http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf`
- LaTeX users:
  - Consider directly generating PDF files using `pdflatex` (especially if you are a MiKTeX user). PDF figures must be substituted for EPS figures, however.
  - Otherwise, please generate your PostScript and PDF files with the following commands:

    ```
    dvips mypaper.dvi -t letter -Ppdf -G0 -o mypaper.ps
    ps2pdf mypaper.ps mypaper.pdf
    ```

    Check that the PDF files only contains Type 1 fonts.
  - xfig "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
  - The `\bbold` package almost always uses bitmap fonts. You can try the equivalent AMS Fonts with command

    ```
    \usepackage[psamsfonts]{amssymb}
    ```

    or use the following workaround for reals, natural and complex:

    ```
    \newcommand{\RR}{I\!\!R} %real numbers
    \newcommand{\Nat}{I\!\!N} %natural numbers
    \newcommand{\CC}{I\!\!\!\!C} %complex numbers
    ```

  - Sometimes the problematic fonts are used in figures included in LaTeX files. The ghostscript program `eps2eps` is the simplest way to clean such figures. For black and white figures, slightly better results can be achieved with program `potrace`.

- MSWord and Windows users (via PDF file):
  - Install the Microsoft Save as PDF Office 2007 Add-in from `http://www.microsoft.com/downloads/details.aspx?displaylang=en&familyid=4d951911-3e7e-4ae6-b059-a2e79ed87041`
  - Select "Save or Publish to PDF" from the Office or File menu
- MSWord and Mac OS X users (via PDF file):
  - From the print menu, click the PDF drop-down box, and select "Save as PDF..."
- MSWord and Windows users (via PS file):
  - To create a new printer on your computer, install the AdobePS printer driver and the Adobe Distiller PPD file from `http://www.adobe.com/support/downloads/detail.jsp?ftpID=204` *Note:* You must reboot your PC after installing the AdobePS driver for it to take effect.
  - To produce the ps file, select "Print" from the MS app, choose the installed AdobePS printer, click on "Properties", click on "Advanced."
  - Set "TrueType Font" to be "Download as Softfont"
  - Open the "PostScript Options" folder
  - Select "PostScript Output Option" to be "Optimize for Portability"
  - Select "TrueType Font Download Option" to be "Outline"
  - Select "Send PostScript Error Handler" to be "No"
  - Click "OK" three times, print your file.
  - Now, use Adobe Acrobat Distiller or ps2pdf to create a PDF file from the PS file. In Acrobat, check the option "Embed all fonts" if applicable.

If your file contains Type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

10.1 Margins in LaTeX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the graphicx package. Always specify the figure width as a multiple of the line width as in the example below using .eps graphics

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.eps}
```

or

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

for .pdf graphics. See section 4.4 in the graphics bundle documentation (`http://www.ctan.org/tex-archive/macros/latex/required/graphics/grfguide.ps`)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command.