Exploring Deep Learning-Based Feature Extraction for Case-Based Reasoning Retrieval

Zachary Wilkerson

Indiana University, Bloomington, IN, United States

Abstract

Case-based reasoning performance is critically dependent on accurate retrieval, which in turn is supported by effective case indexing. Useful indices may be unavailable and/or difficult to generate manually, and so generating features for case indexing using deep learning is an appealing solution. This paper outlines a research plan investigating how deep learning systems may be leveraged and modified to generate high-quality features for case-based retrieval, including a methodology that explores various deep learning models, feature extraction approaches, and training considerations. It also outlines how already-published results make progress in these explorations and how future work will continue to expand upon the existing experimental foundation.

Keywords

Case-Based Reasoning, Deep Learning, Feature Learning, Hybrid Systems, Indexing, Retrieval

1. Introduction

Accurate case-based reasoning (CBR) performance derives from retrieving useful cases from the case base, and effective retrieval depends on the quality of indices used to differentiate cases. Such indices may be defined through a combination of manual knowledge engineering (e.g., [1, 2]) and situation assessment. In these instances, the resulting indexing structures may both capture key aspects of the domain and facilitate convincing explanations for humans. However, manual knowledge engineering can be costly, and indices may be inaccurate or incomplete for poorly understood domains or domains that lack well-defined, human-understandable features (e.g., computer vision tasks).

Methods for addressing these deficiencies have been explored in the literature through symbolic approaches (e.g., [3]), but recent advances in deep learning (DL) make neural models appealing to extract useful information from raw data for CBR. The resulting DL-CBR hybrid systems ideally combine the inherent interpretability of CBR models via case presentation with the inference/learning power of DL models. Various combinations of the two types of systems have been researched, such as “twin systems" that leverage extracted weights from a network model to guide CBR retrieval of explanatory cases [4] and injecting CBR knowledge/structure into DL “prototype network" systems to increase their interpretability (e.g., [5]). In particular, hybrid systems leveraging DL models to extract features for CBR retrieval appear especially promising (e.g., [6, 7]).

ICCBR DC'23: Doctoral Consortium at ICCBR2023, July 17 – 20, 2023, Aberdeen, Scotland
zachwilk@indiana.edu (Z. Wilkerson)

However, despite the growing call to address the relative opacity of end-to-end DL systems, and even with the relative success of DL-CBR hybridization, the numerous variables that influence DL-based feature extraction for CBR are relatively unexplored at present. This research aims to investigate this space to better understand how to optimize DL-CBR hybridization for feature generation and similarity assessment for the highest possible accuracy.

2. Research Plan

This research broadly investigates different variables/approaches that inform a general DL model for feature extraction for CBR retrieval, with the overall goal of maximizing CBR retrieval accuracy. Such extracted features may replace or supplement features developed through manual knowledge engineering, depending on the domain.

2.1. Research Objectives

Under these broad investigation goals, this research aims to address three primary objectives:

1. **Quantify how using extracted features and knowledge-engineered features in concert may increase retrieval accuracy.** In applications where some domain knowledge is already present, extracted features may supplement existing features. Thus, it is important to evaluate the circumstances (e.g., accuracy/comprehensiveness of knowledge-engineered features in the domain) under which using both feature sets together is most effective. Such an evaluation may be potentially influenced by a "curse of dimensionality" resulting from large numbers of extracted features, independence (or lack thereof) between knowledge-engineered and extracted features, and/or the continuous nature of extracted features versus generally discrete knowledge-engineered ones.
2. **Investigate the impact of network structures on extracted feature quality.** Neural architectures are a thoroughly-researched aspect of DL literature, and just as structure significantly influences model performance and domain applicability, it may have a similarly significant impact on CBR feature quality. As a necessarily incomplete list of examples, the location of feature extraction within the network may influence the complexity of extracted features, the DL model used may influence the way in which features capture domain information (indeed some models may generate features that are most applicable to certain domains/problems), and modifications to existing DL structures (e.g., the addition/modification of layers and/or their properties) may further influence how features are generated by the DL model.
3. **Explore interplay between DL and CBR needs during network training.** Previous research on DL feature extraction for CBR retrieval operates on the plausible assumption that useful features for end-to-end DL classification are also useful for CBR retrieval. However, given that DL and CBR systems have different classification methodologies, require different amounts of training data, etc., this may not always be the case. Exploring these variables alongside developing novel implementations that focus on the coupling of DL and CBR models (e.g., using CBR performance as a loss component to guide DL

backpropagation, and exploring methods for doing so with minimal concessions for training efficiency) may further increase feature quality.

Furthermore, these objectives do not operate in a vacuum. For one, the existence of knowledge-engineered and/or extracted weights may further influence feature extraction; indeed, DL-based weight extraction may exhibit similar patterns as feature extraction or derive from an independent set of variables. For another, while case presentation is a useful medium for explanation, the lack of interpretability of features extracted from DL models may limit its effectiveness; additional per-feature explanation (e.g., [8]) and/or explanation-oriented retrieval strategies (e.g., [9]) may be applicable to augment DL-CBR interpretability. Such projects arguably extend beyond the doctoral research scope for this work, but they exist as potential future research avenues and important contextual considerations for these experiments.

2.2. Methodology

The research objectives described above are interdependent in their influence on DL feature extraction for CBR retrieval, a fact that is underscored by current research progress to date [10, 11, 12]. My previously completed experiments have established proof-of-concept implementations that highlight the accuracy benefits of using both knowledge-engineered and extracted features in concert [10], have evaluated how extracting features from different locations in the DL model and extracting different numbers of features impact feature quality based on retrieval accuracy [11], and have explored how using different DL models for feature extraction (i.e., VGGNet, Inception V3, and DenseNet) influence feature quality based on retrieval accuracy [12].

Looking ahead, the research plan going forward will continue on this foundation, with special focus on maximizing CBR retrieval accuracy. Specifically, investigations such as integrating CBR loss into backpropagation during training and using pretrained DL systems to minimize the amount of training data required to establish a high-quality feature extraction model will focus on balancing DL and CBR needs in the feature extraction process. These will be supplemented by experimental results from finer-grained investigations on the horizon, such as exploring other models (e.g., MLP mixers and transformers) as feature extractors, using case-based maintenance methods to minimize case base size while still allowing for larger training data sets to increase extracted feature quality, and/or evaluating a refined DL-CBR model(s) for a variety of image classification domains (e.g., ImageNet, MNIST, etc.).

3. Progress Summary

Research to date has focused on establishing a conceptual foundation and proof-of-concept-level implementations for DL-CBR hybridization for retrieval. The pertinent experiments have supported a general architecture for DL-based feature extraction supporting CBR retrieval (Figure 1) and have established the following specific conclusions in their respective publications:

1. **Using extracted and knowledge-engineered features in concert can increase retrieval accuracy [10].** Especially in instances where knowledge-engineered features are useful but incomplete, supplementing the existing feature set with extracted features can produce significant accuracy benefits.

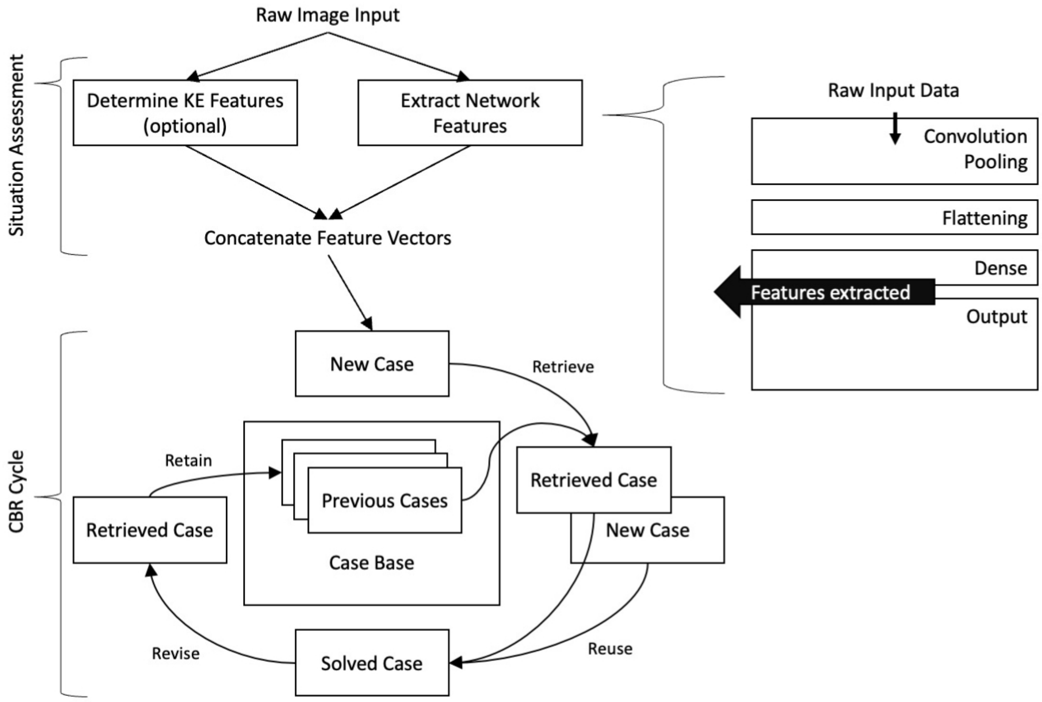

Figure 1: Model diagram for DL-CBR hybrid approach for feature extraction for retrieval (first published in Case-Based Reasoning Research and Development, ICCBR 2023 by Springer Nature [12]).

2. **Extracting features from later in the DL model frequently results in higher-quality features [11].** In particular, while features have typically been extracted immediately following convolution, ideally to capture the atomic elements of an image (e.g., [6, 7]), it appears that extracting features immediately before the output layer (i.e., following the densely-connected layers) of a DL model produces higher-quality features.
3. **There exists a "happy medium" for the number features to extract to maximize feature quality [11].** While DL models require a minimal number of trainable nodes to converge on a solution, CBR systems can incur a "curse of dimensionality" given too many features. As a result, optimal retrieval accuracy can be achieved at a "happy medium" between these competing needs.
4. **Feature quality derives strongly from a trade-off between model complexity and training data requirements [12].** In general, there appears to be no "catchall" model for feature generation, especially in the data-sparse scenarios in which CBR systems are frequently applied. Thus, more complex models are not necessarily better than simpler models that can generalize better from less training data.

Taken together, these conclusions support a broad foundation for DL-CBR hybridization for retrieval that has the potential to guide future research applying DL principles to CBR. Future work will focus on questions for further optimizations, such as integrating CBR needs into the DL training process and maintaining a reasonable case base size while allowing for larger training data sets to support the DL system, among other refinements and explorations.

References

[1] D. Leake, An indexing vocabulary for case-based explanation, in: Proceedings of the Ninth National Conference on Artificial Intelligence, AAAI Press, Menlo Park, CA, 1991, pp. 10–15.

[2] R. Schank, M. Brand, R. Burke, E. Domeshek, D. Edelson, W. Ferguson, M. Freed, M. Jona, B. Krulwich, E. Ohmayo, R. Osgood, L. Pryor, Towards a general content theory of indices, in: Proceedings of the 1990 AAAI spring symposium on Case-Based Reasoning, AAAI Press, Menlo Park, CA, 1990.

[3] S. Bhatta, A. Goel, Model-based learning of structural indices to design cases, in: Proceedings of the IJCAI-93 Workshop on Reuse of Design, IJCAI, Chambery, France, 1993, pp. A1–A13.

[4] E. M. Kenny, M. T. Keane, Twin-systems to explain artificial neural networks using case-based reasoning: Comparative tests of feature-weighting methods in ANN-CBR twins for XAI, in: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, 2019.

[5] C. Chen, O. Li, D. Tao, A. Barnett, C. Rudin, J. K. Su, This looks like that: Deep learning for interpretable image recognition, in: Advances in Neural Information Processing Systems 32, Curran, 2019, pp. 8930–8941.

[6] S. Sani, N. Wiratunga, S. Massie, Learning deep features for kNN-based human activity recognition, in: Proceedings of ICCBR 2017 Workshops (CAW, CBRDL, PO-CBR), Doctoral Consortium, and Competitions co-located with the 25th International Conference on Case-Based Reasoning (ICCBR 2017), Trondheim, Norway, June 26-28, 2017, volume 2028 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2017, pp. 95–103.

[7] J. T. Turner, M. W. Floyd, K. M. Gupta, T. Oates, NOD-CC: A hybrid CBR-CNN architecture for novel object discovery, in: Case-Based Reasoning Research and Development, ICCBR 2019, 2019, pp. 373–387.

[8] K. Bach, P. J. Mork, On the explanation of similarity for developing and deploying cbr systems, in: FLAIRS, 2021.

[9] D. Doyle, P. Cunningham, D. Bridge, Y. Rahman, Explanation oriented retrieval, in: Advances in Case-Based Reasoning, ECCBR 2004, volume 3155, 2004, pp. 157–168.

[10] Z. Wilkerson, D. Leake, D. Crandall, On combining knowledge-engineered and network-extracted features for retrieval, in: Case-Based Reasoning Research and Development, ICCBR 2021, 2021, pp. 248–262.

[11] D. Leake, Z. Wilkerson, D. Crandall, Extracting case indices from convolutional neural networks: A comparative study, in: Case-Based Reasoning Research and Development, ICCBR 2022, 2022, pp. 81–95.

[12] D. Leake, Z. Wilkerson, D. Crandall, Examining the impact of network architecture on extracted feature quality for CBR, in: Case-Based Reasoning Research and Development, ICCBR 2023 (in print), 2023.