Universitatea Tehnică a Moldovei

FEW-SHOT AUTOMATIC DEEP LEARNING

ÎNVĂȚAREA AUTOMATIZATĂ CU UN NUMĂR MIC DE DATE

Student: Burlacu Alexandru

Conducător: Nistor Grozavu,

conf. univ., dr., HdR

Chişinău, 2020

Contents

DEFINITIONS AND ABBREVIATIONS 5
LIST OF TABLES 8
LIST OF FIGURES 9
1 PROBLEM ANALYSIS 11
1.1 Premises 11
1.2 Scope 12
1.3 Requirements 13
1.4 Conclusion 14
2 EXISTING SOLUTIONS 15
2.1 Content retrieval 16
2.2 Self-supervised learning 17
2.3 Few-shot learning 18
2.4 Hyperparameter optimization 19
2.5 Conclusion 20
3 IMPLEMENTATION 22
3.1 Dataset enrichment 23
3.2 Self-supervised component 23
3.3 Few-shot learning component 25
3.4 Hyperparameter optimization loop 27
3.5 Experiment findings 29
3.6 The API 29
3.7 Conclusion 34
CONCLUSIONS AND FUTURE WORK 35

ABSTRACT

Machine learning has proved to be a powerful tool in solving numerous real-world problems. From assisting in cancer and diabetes diagnosis to detecting frauds, spam emails and even driving cars and landing rockets, deep neural networks can be used everywhere. This is the reason so many people believe and want to try attacking their problems using deep learning algorithms, but frequently this is a non-trivial task due to the lack of a big enough dataset. Deep learning is especially known for high data appetites to work well.

This work tries to reduce the need for annotated data to design and train reasonably high performing neural networks for a given problem by using so-called few-shot learning, in combination with other techniques to further reduce the need for a big, diverse, and preferably well-designed dataset, thus extending the range of use cases for deep learning techniques. Such models could also be used to help in annotation efforts.

The primary assumptions of this project are, (1) it is possible to use few-shot learning techniques to obtain performant models on small annotated datasets, (2) it is possible to reduce the need for annotated data via self-supervision and by enriching the original dataset with datasets approximately similar with it.

The motivation and the analysis of the system requirements are further described in the first chapter of this work, the existing solutions and the varied research this work draws from is summarised in the second chapter, while the details about the algorithms for all stages of the system can be found in the third chapter.

Chapter four and five describe the system's design and implementation, and the evaluation procedure of the system. Finally, the thesis ends with a conclusion and discussion section.

REZUMAT

Învățarea automată s-a dovedit a fi un instrument puternic în rezolvarea a numeroase probleme din lumea reală. De la asistență în diagnosticarea cancerului și diabetului până la depistarea fraudelor, e-mailurilor spam și chiar conducerea mașinilor și aterizarea rachetelor, rețelele neuronale pot fi folosite peste tot. Acesta este motivul pentru care mulți oameni cred și vor să încerce să rezolve probleme folosind algoritmi de învățare aprofundată, dar frecvent aceasta este o sarcină netrivială din cauza lipsei unui set de date suficient de mare. Învățarea aprofundată este cunoscută mai ales pentru apetitul mare pentru date pentru a funcționa bine.

Această lucrare încearcă să reducă nevoia de date adnotate pentru a proiecta și a antrena rețele neuronale cu performanțe rezonabile pentru o anumită problemă, utilizând așa-numita învățare cu puține exemple, în combinație cu alte tehnici pentru a reduce în continuare nevoia de un set de date mare, bine conceput, și divers, extinzând astfel gama de cazuri de utilizare pentru tehnici de învățare aprofundată. Astfel de modele ar putea fi folosite și pentru a ajuta la eforturile de adnotare.

Ipotezele principale ale acestui proiect sunt: (1) este posibil să se utilizeze tehnici de învățare cu puține exemple pentru a obține modele performante pe seturi mici de date adnotate, (2) este posibil să se reducă nevoia de date adnotate prin învățare cu auto-supervizare și prin îmbogățirea setului de date original cu seturi de date aproximativ similare cu acesta.

Motivația și analiza cerințelor pentru sistem sunt descrise în continuare în primul capitol al acestei lucrări, soluțiile existente și cercetările variate din care se inspiră această lucrare sunt rezumate în al doilea capitol, în timp ce detaliile despre algoritmii pentru toate etapele sistemului pot fi pot fi găsite în capitolul al treilea.

Capitolele patru și cinci descriu proiectarea și implementarea sistemului și procedura de evaluare a sistemului. În cele din urmă, teza se încheie cu o secțiune de concluzii și discuții.

INTRODUCTION

The purpose of this entire document is to provide an overview of the research project that aims to develop a set of methods that would allow training deep learning models with small amounts of annotated data.
In addition, this document defines the problem that the project tries to solve, explains the workflow, structural and algorithmic internals of the software, but also summarizes the research area, its developments and the existing sub-domains.

Analyzing the requirements for an easy-to-use system with superior performance in different areas and problems, we concluded that to meet these requirements, a composite, systemic approach is needed, combining different existing methods, and possibly developing new solutions.

This project's aim is to develop a system capable of learning an efficient, production-ready deep learning model while minimizing the amount of data needed to obtain it. For this, we set out to combine several techniques that in combination we hope will have a synergistic effect.

The system is supposed to be a tool that will be primarily focused on fast prototyping, for either quick exploratory projects or as a seed model through which better, more optimized deep learning models can be trained, even on more data if the seed model is used as an annotation tool. It's primary target is to democratize development of neural networks for domains with limited data and/or with low quality data.
Born out of frustration, or otherwise limitations of existing solutions, the system being based on most recent tools for FSL and self-supervised learning can significantly reduce the costs for development of deep neural networks, especially so for domains with little amounts of annotated data. Through hyperparameter optimization, the system can create a high-performance and general system that could facilitate rapid exploratory projects and subsequent accumulation of annotated data through active learning methods.

The thesis is organized as follows: first, in the Domain Analysis chapter, the problem that the project tries to solve will be defined and explained, also the requirements and primary assumptions will be outlined; the Existing Solutions chapter is second, it summarises the varied research and tools this work is build from; then in the third chapter, System Analysis and Design the solution will be described from the perspective of use cases, workflow and components, also diving into explaining the implementation details, the tools and algorithms used; chapter four and five describe the system's design and implementation, and the evaluation procedure of the system respectively. Finally, the thesis ends with a conclusion and discussion section.

References

[1] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural net-works,”Communications of the ACM, 2012.

[2] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao,K. Macherey, J. Klingner, A. Shah, M. Johnson, X. Liu, Łukasz Kaiser, S. Gouws, Y. Kato, T. Kudo,H. Kazawa, K. Stevens, G. Kurian, N. Patil, W. Wang, C. Young, J. Smith, J. Riesa, A. Rudnick,O. Vinyals, G. Corrado, M. Hughes, and J. Dean, “Google’s neural machine translation system: Bridging the gap between human and machine translation,”CoRR, vol. abs/1609.08144, 2016.

[3] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255, 2009.

[4] A. Burlacu, “Overview of computer vision supervised learning techniques for low-data training,” 2019.

[5] A. Burlacu, “Bringing together few-shot learning and self-supervision for greater sample efficiency,”2020.

[6] Y. Wang, Q. Yao, J. T. Kwok, and L. Ni, “Generalizing from a few examples,”ACM Computing Surveys(CSUR), vol. 53, pp. 1 – 34, 2020.

[7] J.-C. Su, S. Maji, and B. Hariharan, “When does self-supervision improve few-shot learning?,”ArXiv,vol. abs/1910.03560, 2019.

[8] Z. Ji, X. Zou, T. Huang, and S. Wu, “Unsupervised few-shot feature learning via self-supervised training,”Frontiers in Computational Neuroscience, vol. 14, p. 83, 2020.

[9] J. J. Lim, R. Salakhutdinov, and A. Torralba, “Transfer learning by borrowing examples for multiclass object detection,” inProceedings of the 24th International Conference on Neural Information ProcessingSystems, NIPS’11, (Red Hook, NY, USA), p. 118–126, Curran Associates Inc., 2011.

[10] Ming-Kuei Hu, “Visual pattern recognition by moment invariants,”IRE Transactions on InformationTheory, vol. 8, no. 2, pp. 179–187, 1962.

[11] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in2005 IEEE ComputerSociety Conference on Computer Vision and Pattern Recognition (CVPR’05), vol. 1, pp. 886–893 vol. 1,2005.

[12] Scikit-Image, “Gabors / primary visual cortex “simple cells” from an image.”

[13] C. D. Manning, P. Raghavan, and H. Sch ̈utze,Introduction to Information Retrieval. USA: Cambridge University Press, 2008.

[14] J. E. Moody, “Fast learning in multi-resolution hierarchies,” inAdvances in Neural Information Processing Systems 1, [NIPS Conference, Denver, Colorado, USA, 1988](D. S. Touretzky, ed.), pp. 29–39,Morgan Kaufmann, 1988.

[15] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,”ArXiv, vol. abs/1310.4546, 2013.

[16] J.-T. Huang, A. Sharma, S. Sun, L. Xia, D. Zhang, P. Pronin, J. Padmanabhan, G. Ottaviano, and L. Yang,“Embedding-based retrieval in facebook search,”Proceedings of the 26th ACM SIGKDD InternationalConference on Knowledge Discovery Data Mining, 2020.

[17] M. Meenalochini, K. Saranya, G. V. Rajkumar, and A. Mahto, “Perceptual hashing for content based im-age retrieval,” in2018 3rd International Conference on Communication and Electronics Systems (ICCES),pp. 235–238, 2018.
[18] T. Chen, S. Kornblith, M. Norouzi, and G. E. Hinton, “A simple framework for contrastive learning of visual representations,”ArXiv, vol. abs/2002.05709, 2020.
[19] J.-B. Grill, F. Strub, F. Altch ́e, C. Tallec, P. H. Richemond, E. Buchatskaya, C. Doersch, B. A. Pires, Z. D.Guo, M. G. Azar, B. Piot, K. Kavukcuoglu, R. Munos, and M. Valko, “Bootstrap your own latent: A new approach to self-supervised learning,”ArXiv, vol. abs/2006.07733, 2020.
[20] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,”arXiv preprint arXiv:1810.04805, 2018.
[21] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” 2019.
[22] A. Zisserman, “Self-supervised learning,” 2019.
[23] P. Agrawal, J. Carreira, and J. Malik, “Learning to see by moving,”2015 IEEE International Conference on Computer Vision (ICCV), Dec 2015.
[24] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, “Deep clustering for unsupervised learning of visual features,” inECCV, 2018.
[25] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sas-try, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M.Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner,S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, “Language models are few-shot learners,” 2020.
[26] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,”ArXiv, vol. abs/1703.03400, 2017.
[27] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, “Matching networks for one shot learning,”ArXiv, vol. abs/1606.04080, 2016.
[28] J. Snell, K. Swersky, and R. Zemel, “Prototypical networks for few-shot learning,”ArXiv,vol. abs/1703.05175, 2017.
[29] A. Lacoste, H. Larochelle, F. Laviolette, and M. Marchand, “Sequential model-based ensemble optimiza-tion,”Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014, 02 2014.
[30] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, “Hyperband: A novel bandit-based approach to hyperparameter optimization,”J. Mach. Learn. Res., vol. 18, pp. 185:1–185:52, 2017.
[31] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,”J. Mach. Learn. Res.,vol. 13, p. 281–305, Feb. 2012.
[32] N. Hansen,The CMA Evolution Strategy: A Comparing Review, vol. 192, pp. 75–102. 06 2007.
[33] A. Coates, H. Lee, and A. Ng, “An analysis of single layer networks in unsupervised feature learning,”2011.

[34] Y. M. Asano, C. Rupprecht, and A. Vedaldi, "A critical analysis of self-supervision, or what we can learn from a single image,"arXiv: Computer Vision and Pattern Recognition, 2020.
[35] B. Sch¨olkopf, A. Smola, and K.-R. M¨uller, "Nonlinear component analysis as a kernel eigenvalue prob-lem," 1996.
[36] A. Chaudhary, "A visual exploration of deepcluster," 2020 .https://amitness.com/2020/04/illustrated-deepcluster.
[37] R. Liaw, E. Liang, R. Nishihara, P. Moritz, J. E. Gonzalez, and I. Stoica, "Tune: A research platform for distributed model selection and training,"arXiv preprint arXiv:1807.05118, 2018.
[38] M. Jaderberg, V. Dalibard, S. Osindero, W. Czarnecki, J. Donahue, A. Razavi, O. Vinyals, T. Green,I. Dunning, K. Simonyan, C. Fernando, and K. Kavukcuoglu, "Population based training of neural net-works,"ArXiv, vol. abs/1711.09846, 2017.
[39] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein,L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner,L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library,"inAdvances in Neural Information Processing Systems 32(H. Wallach, H. Larochelle, A. Beygelzimer,F. d'Alch´e-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019.
[40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer,R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay,"Scikit-learn: Machine learning in Python,"Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[41] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "Autoaugment: Learning augmentation policies from data," 2018.