The 4th Annual Student Research Project Conference

Stiftung Universität Hildesheim
December 3, 2020

Contents

1. **Session 1** (Chair: Rafael Rego Drumond) **3**
2. **Session 2** (Chair: Mofassir ul Islam Arif) **6**
3. **Session 3** (Chair: Ahmed Rasheed) **9**
4. **Session 4** (Chair: Kiran Madhusudhanan) **12**

SESSION 1

Chair: Rafael Rego Drumond

Model- Agnostic Meta-Learning in the Context of Speech Classification

Team: Ece Atalay, Josué D.Rodriguez Quintana, Naveen kumar Ramu, Santiago Tena Hernandez
Supervisor: Rafael Rego Drumond

Abstract

Deep Learning based speech recognition algorithms have become ubiquitous. They require vast amounts of data for training purposes. Model Agnostic Meta Learning is a candidate solution for training new models using less data. Our tests show that the inheritance from previously performed similar tasks lead to a superior data use efficiency and a superior accuray in comparison to random initialization approaches.

References

- Warden, Pete. "Speech commands: A dataset for limited-vocabulary speech recognition." *arXiv preprint arXiv:1804.03209* (2018).
- Paliwal, Kuldip K. "Decorrelated and liftered filter-bank energies for robust speech recognition." *Sixth European Conference on Speech Communication and Technology*. 1999.
- Zhang, Yundong, et al. "Hello edge: Keyword spotting on microcontrollers." *arXiv preprint arXiv:1711.07128* (2017).

Speaker Identification with Noise Filtering

Team: Yuvaraj Prem Kumar, Shobhit Agarwal, Souaybou Bagayoko, Alina Nicoleta Dobre
Supervisor: Rafael Rego Drumond

Abstract

Speaker identification is an essential task in the field of security and speech processing. In the literature, we find many works that are able to identify a speaker in clear audio speech conditions fully. However, it is not feasible to consider clear audio data in a real-life situation where audio is usually distorted by background noise or device interference. In this research, the goal is to be able to filter the audio, or it's extracted features to facilitate speaker identification. Such techniques might include masking, de-noising auto-encoders, and generative adversarial networks. You must understand the most recent techniques and theory behind it and be able to improve the performance of the state-of-art SID models.

References

- May, Tobias. "Influence of binary mask estimation errors on robust speaker identification." Speech Communication 87 (2017): 40-48.
- May, Tobias, and Timo Gerkmann. "Generalization of supervised learning for binary mask estimation." Acoustic Signal Enhancement (IWAENC), 2014 14th International Workshop on. IEEE, 2014.
- Tirumala, Sreenivas Sremath, and Seyed Reza Shahamiri. "A review on Deep Learning approaches in Speaker Identification." Proceedings of the 8th international conference on signal processing systems. ACM, 2016.

Latent Embedding Optimization for Few-shot

Team: Onno Niemann, Temiloluwa Adeoti, Istiaque Mannafee,
Gopika Sudhakaran, Jennifer Deborah
Supervisor: Lukas Brinkmeyer

Abstract

Few-shot classification has become an immensely popular problem in the deep learning community. It deals with the classfication of images of a new task with novel labels after observing only a handful of samples of the respective task. Various meta-learning approaches drastically improved the state-of-the-art in recent years. However, only a few approaches focus on few-shot segmentation. The task is to analyze current state-of-the-art approaches for few-shot classification and evaluate them for their applicability to segmentation tasks while innovating additional strategies to adapt these methods more closely to a segmentation problem.

References

- Finn, Chelsea, Pieter Abbeel, and Sergey Levine. "Model-agnostic meta-learning for fast adaptation of deep networks." Proceedings of the 34th International Conference on Machine Learning-Volume 70. MLR. org, 2017.
- A. A. Rusu, D. Rao, J. Sygnowski, O. Vinyals, R. Pascanu, S. Osindero, and R. Hadsell. Meta-learning with latent embedding optimization. In International Conference on Learning Representations, 2019.
- Rakelly, Kate, et al. "Few-shot segmentation propagation with guided networks." arXiv preprint arXiv:1806.07373 (2018).

SESSION 2

Chair: Mofassir ul Islam Arif

Personalized Recommendation of Parking Places

Team: Jung Eun Park, Surya Jeevanantham, Lavanya Chinthakayala, Boma Tonye Dagogo, Sakshi Singh
Supervisor: Eya Boumaiza

Abstract

The issue of parking vehicles in over-populated cities is becoming a concern and needs attention. Traditional recommendation systems contemplate a static set of elements and attempt to recommend pertinent items to the users. This research project serves as the novel technique of recommendation systems and provide personalized parking slots to individual users based on their historical data. We build a Context-Aware recommender system which exploits the contextual information from the user. Comparison between the traditional Context-aware RS and Deep-learning based RS is performed, concluding that the latter outperforms with time as the contextual feature. We show that considering time at which the user perfer's to park acts as an additional information and hence the model outperforms all the existing baseline models.

References

- Yang S., Ma W., Pi X., Sean Q, (2019) A deep learning approach to real-time parking occupancy prediction in spatio-temporal networks incorporating multiple spatio-temporal data sources [Yang S. et al., 2019]
- Context-Aware Recommender Systems in Mobile Scenarios [Woerndl W. et al.,2009]
- Context-aware recommender system: A review of recent developmental process and future research direction [Haruna K. et al., 2017]
- Time-aware recommender systems: a comprehensive survey[Campos et al., 2014]
- Personalized Recommendation Using Time Series Analysis [Zhang et al., 2015]

Learning Representation for Active Learning with Latent Spaces

Team: Zain Ulabidin, Luis Diego Rosello, Juan Fernando Espinosa, Fizza Batool
Supervisor: Daniel Pototzky & Hadi Samer Jomaa

Abstract

Sophisticated data augmentation strategies have shown to significantly boost object detection performance. However, little research has been done on using data augmentation for solving common problems like crowd scenes, heavy occlusion and distinguishing similar classes like cows and horses. Expanding upon one of the existing methods (like a bag of freebies) might improve performance on those tasks.

References

- Barret Zoph, Ekin D. Cubuk, Golnaz Ghiasi, Tsung-Yi Lin, Jonathon Shlens, Quoc V. Le (2019): Learning Data Augmentation Strategies for Object Detection. https://arxiv.org/abs/1906.11172
- Zhi Zhang, Tong He, Hang Zhang, Zhongyue Zhang, Junyuan Xie, Mu Li (2019): Bag of Freebies for Training Object Detection Neural Networks. https://arxiv.org/abs/1902.04103
- Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, Yi Yang (2017): Random Erasing Data Augmentation. https://arxiv.org/abs/1708.04896

Regularization: Extending Disturb Label Method

Team: Hanna Lukashonak, Paweena Tarepakdee, Yongho Kim, Klavdiia Zavalich
Supervisor: Mofassir ul Islam Arif

Abstract

This work extends one of the recently proposed methods of regularization for neural networks – Disturb Label [Xie et al.2016], which was applicable only to classification task. This similar technique is extended in our work onto regression problems demonstrating its efficiency, beating baseline meth- ods for 8 tested datasets, and cooperate well with other regularization methods. We also propose more efficient way to apply Disturb Label method to clas- sification tasks by disturbing selectively only on the confident predictions. This extension helped to im- prove misclassification rate for 5 out of 6 datasets compared to vanilla Disturb Label.

References

- Xie, Lingxi, et al. "Disturblabel: Regularizing cnn on the loss layer." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

SESSION 3

Chair: Ahmed Rasheed

Large Scale Text Classification with Stochastic TextGCN

Team: Eunju Park, Saragol Barakchian, Hyeyeon Kim, Olalekan Olayemi, Hossein Pourtavakoli
Supervisor: Thorben Werner

Abstract

Graph Convolutional Network has been great choice for text classification because of its outstanding performance, but suffers from high memory consumption as a result of building/training the whole corpus graph at once. We tried uniform sampling in layer level but noticed neighborhood explosion problems. Also, we found that reducing graph size by building several smaller graphs instead of a single large graph did not have performance improvement. To address all these aforementioned issues, we proposed a dense graph approach with small datasets and Graph-level sampling approach with large datasets.

References

- Yao, Liang, Chengsheng Mao, and Yuan Luo. "Graph convolutional networks for text classification." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 2019.

Speech Synthesis and Translation via Exemplar Autoenconders

Team: Ajith Gumudavelly, Can Shenol Berk, Christian Sadiki Tyty, Haruki Honda
Supervisor: Shayan Jaweed

Abstract

-

References

- -

Dynamic Graph Embedding

Team: Ahmed Hafez, Atulya Praphul, Youssef Jaradat, Ezeani Godwin
Supervisor: Ahmed Rasheed

Abstract

Graph embedding has been a better alternative to graph analytics methods due its less computational and space cost. However, the graph embedding baseline methods, such as deep walk and singular value decomposition, face a problem dealing with the real-world-dynamic networks in which there is evolving and changing over time. Therefore, a new subfield by the name dynamic graph embedding has been launched to deal with the real-life networks, mostly the methods in this subfield rely on previous static network embedding ideas and follow it by a temporal evolution layer to express the evolution over time. We propose ConvDySAT as an enhancement of DYSAT [1], one of the state-of-the-art methods, by augmenting convolution neural network with the self-attention mechanism which was used in DYSAT to express the structural and temporal evolution. Experimental results on multiple real-world dynamic graph datasets show significant performance gains for ConvDySAT over various state-of-the-art methods.

References

- Sankar, Aravind, et al. "DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks." *Proceedings of the 13th International Conference on Web Search and Data Mining*. 2020.
- Veličković, Petar, et al. "Graph attention networks." *arXiv preprint arXiv:1710.10903* (2017).
- Li, Shiyang, et al. "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting." *Advances in Neural Information Processing Systems*. 2019.

SESSION 4

Chair: Kiran Madhusudhanan

Improving the Accuracy of Visual Explanations Using Gradcam Technique

Team: Srishti Alchoni, Yakub Hafazoglu, Muhammad Hamad Khan, Fahad Fiaz, Fiza Bari
Supervisor: Randolf Scholz

Abstract

Gradient-based Class Activation Mapping or GRAD-CAM develops trust and increases transparency of deep Convolutional Neural Networks by showing why a model predicts what it predicts. As mentioned in the baseline paper, "GRAD-CAM uses the gradients of any target concept flowing into the final convolutional layer to produce a coarse localization map highlighting the important regions in the image for predicting the concept" (Selvaraju, et al, 2017). Once we know where the model is aiming, we want to build our research on top of it by enabling the model to make predictions by only looking at a smaller instance of the object in the image in contrast to the entire available object in the first stage dataset. We continue with the mentioned procedure until the aimed object is completely diminished from the image. This results in a hierarchy of models each trained on different instances of the object in the image. At the end we create an ensemble of these models and test it to find the accuracy. We also claim that our approach with GRAD-CAM can be used to segment objects of interest in an image by finding the threshold*, for which the IOU between the GRAD-CAM mask (mask created to hide the part of the object where the model is looking) and the segmentation mask is the maximum.
* threshold determines the amount of object to be covered by looking at the heatmap obtained from GRADCAM.

References

- Selvaraju, Ramprasaath R., et al. "Grad-cam: Visual explanations from deep networks via gradient-based localization." *Proceedings of the IEEE international conference on computer vision*. 2017.
- Mahmood, Ammar, et al. "Automatic hierarchical classification of kelps using deep residual features." *Sensors* 20.2 (2020): 447.

Reinforcement Learning for Routing Problems

Team: Kirill Tkachuk, Tim Dernedde, Simon Boeder, Ahmad Bdeir
Supervisor: Jonas Falkner

Abstract

Routing problems have for long been very im- portant in business and industry applications. Finding the best routes of delivery vehicles or finding the best arrangement of trading goods in a warehouse are a few examples for such problems. Classical algorithms are often specialized and limited on a single problem and are not applicable to other problems. Recent research in machine learning has shown that routing problems can also be solved by reinforcement learning methods. To do so, most publications used an on-policy leaming method. We propose an off-policy reinforcement learning model called RP- DON, which achieves near state-of-the-art performance on the capacitated vehicle routing problem (CVRP). Furthemore, we extend our model to solve the multi-depot vehicle routing problem (MDVRP). To the best of our knowledge, this is the first reinforcement learning approach to this problem type. Our model is easily adaptable to other routing problems.

References

- Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, and Le Song. Learning Combinatorial Optimization Algorithms over Graphs. In Advances in Neural Information Processing Systems 30, pages 6348–6358. Curran Associates, Inc., 2017.
- Michel Deudon, Pierre Cournut, Alexandre Lacoste, Yossiri Adulyasak, and Louis-Martin Rousseau. Learning Heuristics for the TSP by Policy Gradient. In Willem-Jan van Hoeve, editor, Integration of Constraint Programming, Artificial Intelligence, and Operations Research, volume 10848, pages 170–181. Springer International Publishing, Cham, 2018.
- Kool, Wouter, Herke van Hoof, and Max Welling. "Attention, Learn to Solve Routing Problems!." arXiv preprint arXiv:1803.08475 (2018).

Anomaly Detection: Supervised vs. Unsupervised

Team: Tanvi Joglekar, Ragini Pant, Rustem Devletov, Sara Naz Haider, Talha Abdullah
Supervisor: Kiran Madhusudhanan

Abstract

Anomaly detection is an important aspect in every business process that identifies significant events or changes in the problem domain and affects the predictions. To be able to take advantages of opportunities or fix the data problems, one needs to know what the available data implies, which makes real-time anomaly detection necessary for modern problem domains. The aim of the project different learning is to test and compare approaches for anomaly detection in multivariate time-series data, as well as to investigate how much supervision is required to achieve an the purely accuracy superior to that of unsupervised model. The two real-world NASA datasets, SMAP and MSL, are used for the research.

References

- Su, Ya, et al. "Robust anomaly detection for multivariate time series through stochastic recurrent neural network." *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019.
- Hundman, Kyle, et al. "Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding." *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2018.
- Ruff, Lukas, et al. "Deep semi-supervised anomaly detection." *arXiv preprint arXiv:1906.02694* (2019).