

Visualizing Generative Sum-Product Networks on Image Completion

Proposta de Projeto MAC6914

Renato Lui Geh, NUSP: 8536030

1 Introdução

Redes soma-produto (SPNs, de Sum-Product Networks) são modelos probabilísticos baseados em grafos (PGMs, de Probabilistic Graphical Models) que representam distribuições de probabilidade tratáveis. PGMs clássicas, como Redes Bayesianas e Redes de Markov, têm inferência e aprendizado exato intratável, o que torna inferência aproximada a alternativa factível. Em contrapartida, SPNs possuem inferência exata em tempo linear, além de possuírem uma arquitetura profunda que garantem grande expressividade [2]. Por causa destas características, SPNs tiveram resultados promissoras em várias aplicações, incluindo compleição e classificação de imagem [5, 4, 3].

Uma SPN é um DAG cujos nós internos são somas ponderadas ou produtos. Folhas são distribuições monovariadas tratáveis. O escopo de uma SPN é o conjunto de variáveis que aparecem em sua raiz e todos seus descendentes. É possível interpretar nós somas como relações de semelhança entre os escopos dos filhos, enquanto que produtos podem ser vistos como relações de independência entre filhos. Apesar de inferência exata ser linear, é possível computar mais rapidamente a probabilidade maximum-a-priori (MAP) usando o algoritmo max-product. No entanto, esta alternativa gera resultados aproximados.

A tarefa de compleição de imagem consiste em, dada uma imagem incompleta, prever os valores dos pixels restantes. Este problema é complicado pois muitas

vezes é necessário que o modelo capture tanto interações locais quanto remotas. No caso de interações locais, um exemplo é que valores em uma região próxima tendem a ter cores e intensidade semelhantes. Interações remotas incluem algum objeto que sempre aparece quando outro está na imagem, indicando alguma dependência entre estas regiões.

Neste projeto pretende-se explorar como SPNs gerativas modelam compleição de imagem. Em [6], Vergari *et al* apontam características interessantes de SPNs, inclusive propondo uma interpretação de SPNs como Perceptrons Multi-Camada. No entanto, este projeto busca interpretar SPNs exclusivamente no âmbito de compleição de imagem.

2 Plano de trabalho

A biblioteca GoSPN¹ será usada para construir e aprender as SPNs. Os modelos serão treinados com dois algoritmos de aprendizado para SPNs. O primeiro é baseado na arquitetura de Dennis e Ventura [3], e o segundo no de Gens e Domingos [4].

Serão usados para compleição os datasets Olivetti² e Caltech-101³. O primeiro é um dataset de rostos, contendo 10 imagens de 40 pessoas cada em escala de cinza, chegando num total de 400 imagens. O segundo é um dataset de objetos. São 101 categorias de imagens coloridas de objetos de diferentes tamanhos e resoluções. Cada categoria contém de 40 a até 800 imagens.

Para se entender e visualizar melhor como o modelo representa as dependências e semelhanças entre variáveis, após treinadas as SPNs serão analisadas a cada nó, visualizando-se o escopo que cada filho de um nó representa na imagem. Além disso, espera-se comparar resultados entre compleição de imagem com inferência aproximada, usando o algoritmo max-product, e inferência exata. Finalmente, serão feitos experimentos com o dataset Caltech-101 em escala de cinza e em escala RGB a fim de explorar o quão bem SPNs modelam imagens multicanais com os atuais algoritmos de aprendizado.

¹Disponível em <https://github.com/RenatoGeh/gospn>

²Disponível em <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

³Disponível em http://www.vision.caltech.edu/Image_Datasets/Caltech101/

3 Objetivos

Ao final do artigo, espera-se entender como SPNs treinadas com os algoritmos Dennis-Ventura e Gens-Domingos modelam imagens. Em [1], já foi mostrado que computar o MAP em SPNs é NP-difícil, e adicionalmente que o algoritmo max-product acha aproximações dentro de um fator 2^{c-n} para uma constante $c < 1$. No entanto, quer-se saber se, em compleição de imagem, tal diferença é impactante na prática.

Referências

- [1] Diarmaid Conaty, Denis D. Mauá e Cassio P. de Campos. “Approximation Complexity of Maximum A Posteriori Inference in Sum-Product Networks”. Em: *Uncertainty in Artificial Intelligence 2017 (UAI 2017)* (2017).
- [2] Olivier Delalleau e Yoshua Bengio. “Shallow vs. Deep Sum-Product Networks”. Em: *Advances in Neural Information Processing Systems 24 (NIPS 2011)* (2011).
- [3] Aaron Dennis e Dan Ventura. “Learning the Architecture of Sum-Product Networks Using Clustering on Variables”. Em: *Advances in Neural Information Processing Systems 25* (2012).
- [4] Robert Gens e Pedro Domingos. “Learning the Structure of Sum-Product Networks”. Em: *International Conference on Machine Learning* 30 (2013).
- [5] Hoifung Poon e Pedro Domingos. “Sum-Product Networks: A New Deep Architecture”. Em: *Uncertainty in Artificial Intelligence* 27 (2011).
- [6] Antonio Vergari, Nicola Di Mauro e Floriana Esposito. “Visualizing and Understanding Sum-Product Networks”. Em: *CoRR* abs/1608.08266 (2016).