Sujet de TER Master 2
Experiments on Deep Tensor Learning for Few Shot Learning and Adversarial Robustness

Descriptif

Organisation : Laboratoire d'Informatique Fondamentale (LIF), Université d'Aix Marseille (AMU). Equipe QARMA

Titre : "Experiments on Deep Tensor Learning for Few Shot Learning and Adversarial Robustness"

Contacts : stephane.ayache@lis-lab.fr.

Sujet de TER

Ce sujet de TER s'inscrit dans le contexte de l'apprentissage automatique pour l'analyse, la modélisation et la prédiction de données structurées.

Ces dernières années ont été marquées par l'essor de méthodes d'apprentissages de représentation à base de réseaux de neurones profonds (deep learning) qui sont particulièrement adaptées au traitement de données structurées. En particulier, de nombreuses applications bénéficient aujourd'hui de ces avancées. Des gains très significatifs des performances ont été obtenus en classification d'images, reconnaissance de la parole, compréhension de textes... L'intérêt de ces approches réside dans leur capacité à construire des représentations adaptées à différentes tâches d'apprentissage.

Les réseaux de neurones profonds sont définis à partir de très nombreux paramètres et doivent être appris à partir de très grands volumes de données, ce qui en limite parfois l'usage, par exemple si l'on ne dispose pas de suffisamment de données. Par ailleurs, lorsque appliqués à des données structurées telles que des données spatio-temporelles, incluant une forte redondance (par ex, les pixels voisins d'une vidéo évoluent de façon très similaire), utiliser des couches cachées totalement connectées est peu pertinent car une quantité importante des calculs pourrait être factorisée. Des travaux récents ont proposé des modèles d'apprentissage profonds de tenseurs (où certaines couches cachées sont remplacées par des tenseurs) avec des contraintes sur le rang des tenseur permettant de réduire le temps de calcul et d'améliorer les performances [1, 2]. Cependant des questions concernant le choix optimal du rang du tenseur et l'intégration d'autres contraintes de régularisations autres que le rang restent ouvertes.

Parallèlement à cela, des résultats récents [3] ont montré que les modèles appris par architectures profondes peuvent être très sensibles à de légères perturbations (invisibles à l'oeil nu) rendant totalement erroné la prédiction obtenue par le réseau. Ces travaux ont mené à plusieurs axes de recherche actuels ciblant soit de nouvelles types d'attaques (dites adversarial) ; soit des stratégies de défenses parmi lesquelles la régularisation, la distillation, ou la prise en compte de données perturbées pendant l'entrainement.

L'objectif de ce travail sera de mener des expérimentations pour évaluer l'intérêt des approches par décomposition tensorielle pour deux situations d'intérêts pour la communauté Deep Learning : (1) l'apprentissage de réseaux avec peu d'exemples ; (2) la robustesse de ce type de réseaux aux attaques adversariales. L'étudiant devra prendre en main un code quasiment prêt pour conduire ces expériences, écrit avec les bibliothèques Keras et Tensorflow. Puis, l'étudiant proposera un protocole d'expérimentation et d'évaluation pour parvenir à conclure sur la question posée.

Références

[1] Dong Yu, Li Deng, Brian Hutchinson, "Tensor Deep Stacking Networks", IEEE Transactions on Pattern Analysis & Machine Intelligence vol. 35 no. , p. 1944-1957, , 2013

[2] Novikov, A.; Podoprikhin, D.; Osokin, A. & Vetrov, D. P. (2015), Tensorizing Neural Networks., in Corinna Cortes; Neil D. Lawrence; Daniel D. Lee; Masashi Sugiyama & Roman Garnett, ed., 'NIPS' , pp. 442-450

[3] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in Proceedings of the 2015 International Conference on Learning Representations. Computational and Biological Learning Society, 2015.