

УДК 004.85

Разработка метода учета алеаторной неопределенности для детекции и классификации объектов на изображении.

Лапенко А.С., Университет ИТМО

Консультант – Ханжина Н. Е., аспирант ФИТиП, Университет ИТМО

Научный руководитель – Фильченков А. А., к.ф.-м.-н., доцент, научный сотрудник ФИТиП, Университет ИТМО

В данной работе предлагается модификация функций потерь нейронных сетей *Focal Loss* и *Smooth L1* для решения задачи совместной детекции и классификации объектов на изображениях с учетом алеаторной неопределенности. Результаты тестирования новых функций показали, что применение их при обучении нейронной сети позволяет улучшить результаты решения данных задач.

Введение.

Проблема автоматической детекции и классификации объектов на изображении не является новой. Данную задачу описывали еще в 1997 году^[1], а с развитием глубокого обучения начался новый виток ее разработки, однако до сих пор существующие лучшие решения обладают недостатками, связанными с точностью. Обученные модели для решения этой проблемы используются в различных сферах, таких как сельское хозяйство, медицина, self-driving и другие. Для всех областей возможно обучать свои модели с использованием различных функций потерь. Однако наиболее эффективные – *Focal Loss*^[2], для решения задачи классификации, и *Smooth L1 Loss*^[3] – для регрессии. В данной работе эти функции были изменены так, чтобы они учитывали алеаторную неопределенность. Данные изменения позволяют добиваться более точной детекции и классификации объектов на изображении.

Основная часть.

Для получения новой функции потерь для задачи классификации был использован вывод для логарифмической функции правдоподобия^[4], адаптированный для *Focal Loss* с помощью теоремы Байеса. Полученная функция была прологарифмирована, и была сделана замена параметра алеаторной неопределенности σ , который затем выводится при обучении. В итоге получилась функция *Bayesian Focal Loss*, которая схожа с исходным *Focal Loss*, но учитывает алеаторную неопределенность, а при значении σ равным нулю сводится к *Focal Loss*.

Для получения новой функции потерь для задачи детекции, которая с точки зрения машинного обучения является задачей регрессии, было использовано два разных распределения, так как исходно *Smooth L1 Loss* имеет точку излома. Для первой части функции было использовано распределение Гаусса^[5], а для второй – Лапласа. Аналогично, как и в задаче классификации, каждая из частей *Bayesian Smooth L1 Loss* свелась к виду аналогичному *Smooth L1 Loss* с добавлением нового параметра алеаторной неопределенности σ для обучения в модели. Так же, как и в предыдущей случае, когда σ становится равным нулю, полученная формула становится тождественно равна исходному *Smooth L1 Loss*.

Полученные функции были реализованы на языке программирования *Python* с использованием фреймворка *tensorflow*. Экспериментальное исследование показало необходимость ввести ограничения на исходные функции. Одна из проблем связана с тем, что если σ становится отрицательным, то вся функция потерь становилась отрицательной.

Эту проблему решили путем введения ограничения на σ строго больше нуля. Вторая сложность – новая функция потерь для регрессии оказалась разрывной, из-за чего во время обучения градиент попадал в область разрыва. Эту проблему решили тем, что добавили константу, зависящую от σ , ко второй части, чтобы в месте разрыва обе части новой *Smooth LI Loss* были равны. Данное сглаживание не нарушило приводимость новой функции потерь для задачи регрессии к исходному виду *Smooth LI Loss*.

Результаты.

Разработанные функции потерь были исследованы и протестированы на наборе данных COCO с использованием архитектуры нейронной сети *RetinaNet-ResNet-50*. Лучший результат на тестовом множестве был получен с применением метода оптимизации Adam.

Базовый результат выбранной нейронной сети с исходными функциями потерь при пороге *Intersection over Union (IoU)* от 0,5 до 0,95 составил 0,305 по мере precision, по мере recall – 0,520. После обучения с использованием разработанных функций потерь результат по мере precision составил 0,311, а по мере recall – 0,531 при том же *IoU*.

Результаты показали, что учет алеаторной неопределенности дал прирост в точности детекции и классификации объектов на изображениях. Дальнейшее повышение точности станет предметом будущих исследований.

Список литературы

- [1] Abu-Mostafa Y. S. Learning from hints in neural networks //Journal of complexity. – 1990. – Т. 6. – №. 2. – С. 192-198.
- [2] Lin T. Y. et al. Focal loss for dense object detection //Proceedings of the IEEE international conference on computer vision. – 2017. – С. 2980-2988.
- [3] Girshick R. Fast r-cnn //Proceedings of the IEEE international conference on computer vision. – 2015. – С. 1440-1448.
- [4] Kendall A., Gal Y. What uncertainties do we need in bayesian deep learning for computer vision? //arXiv preprint arXiv:1703.04977. – 2017.
- [5] Kendall A., Gal Y., Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 7482-7491.

Лапенюк А.С. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись