

АНАЛІЗ ТА ПОРІВНЯННЯ ФУНКЦІЙ ВТРАТ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ СЕГМЕНТАЦІЇ

Ковтуненко А.Р.

Науковий керівник – д.т.н., проф. Машталір В.П.

Харківський національний університет радіоелектроніки, каф. ІНФ,
м. Харків, Україна

e-mail: andrii.kovtunenکو@nure.ua

The work is devoted to analyzing and comparing the use of different loss functions for image and video segmentation tasks depending on problem and input data. The work analyzes the advantages and disadvantages of each of the loss functions selected for analysis, for what purpose and conditions each of the functions is suitable and the possibility of combining the properties of each of them to create new loss functions.

Прихід машинного навчання в задачі комп'ютерного зору істотно спростив розв'язання цих задач і поліпшив, у той час, результати. Наприклад, завдання сегментації відео та розпізнавання дій вирішували за допомогою традиційних методів сегментації зображень: порогові методи; методи на основі областей; методи на основі виділення меж; методи на основі алгоритму водоподілу; методи на основі кластеризації з подальшою постобробкою результатів за допомогою евристик або інших детермінованих методів. З приходом машинного та глибокого навчання, а саме згорткових нейронних мереж, з'явилися такі методи, як U-Net, convolutional encoder-decoder, Mask-RCNN, Yolact [1] та інші, які вже самостійно могли вивчати ознаки, враховувати контекст та просторові зв'язки, з'явилась краща стійкість до мінливості вхідних даних та покращилась продуктивність. Для навчання нейронних мереж необхідна функція втрат (loss function).

Функція втрат – це функція, яка призначена для розрахунку помилки (різниці) між значеннями, що були отримані моделлю, та еталонними значеннями для вхідних даних, які були подані моделі. Завдання навчання моделі – мінімізувати значення функції втрат. Розглянемо функції для сегментації.

Cross-entropy loss – давно відома функція, але яка досі використовується в роботах, де вихідні дані можна представити як ймовірності розподілу між класами, і дає змогу отримувати state of the art результати. Переваги: ефективна з обчислювального погляду і може бути легко оптимізована за допомогою градієнтного спуску; інваріантна до масштабування і зсуву прогнозованих ймовірностей. Недоліки – погано показує себе на незбалансованих даних і чутлива до викидів. Функція сильно штрафує модель за неправильні класифікації і вони можуть перебити об'єкти малочисельного класу, які були правильно класифіковані. Це може призвести до занадто сильного зміщення моделі в бік класів, яких більше, та ігнорування класів,

яких менше. Для вирішення цих проблем використовують варіації: *balanced cross-entropy*, *weighted cross-entropy*, *focal loss*.

Focal loss [2] – дає змогу працювати з високодисперсними наборами даних і зменшує внесок правильних відповідей з високою ймовірністю (*easy examples*), що дає змогу моделі більше реагувати на неправильні відповіді, ніж на правильні. Переваги – функція дозволяє моделі фокусуватися на прикладах, у яких модель помиляється, а не на тих, які вона може впевнено передбачити. Недоліки – додаткові гіперпараметри, які теж потребують налаштування.

Dice loss – призначена для порівняння *dice coefficient* між двома множинами, отриманою маскою сегментації та еталонною розміткою. Головна перевага цієї функції те, що вона допомагає боротися з дисбалансом класів, звертаючи менше уваги до фонових пікселів і більше – до пікселів об'єктів. Недоліки: можливий вибух градієнтів під час оптимізації; на великих зображеннях може довго обчислюватися.

Tversky loss [3] – функція втрат, яка дає змогу налаштовувати, як враховувати FP і FN помилки і являє собою модифіковану *dice loss*. Переваги – так само враховує дисбаланс класів і дає змогу це налаштовувати. Недоліки – нестабільна на початковому етапі навчання.

IoU – поширена метрика для оцінювання якості виявлення об'єктів, яка так само може бути використана як функція помилки для початку сегментації – *IoU loss* (*Jaccard loss*). Переваги *IoU*: легка інтерпретованість; інваріантність до розмірів об'єктів. Недоліки – функція має розриви і може бути недеференційована. У разі повної несхожості еталонного об'єкта й отриманого результату градієнт дорівнює нулю, і оптимізація не буде виконана. Для вирішення цієї проблеми було розроблено *Generalized IoU loss* [4].

Boundary loss [5] – функція, яка оцінює межі об'єкта. Наприклад, попередні функції оцінювали знайдені регіони за площею, але кривизна кордону не може дуже вплинути на таку оцінку, але існують задачі, в яких врахування кордонів важливе. *Boundary loss* враховує відстань між істинним контуром і прогнозованим контуром, а не площу перекриття масок. Переваги: враховує границі об'єктів; кращий поділ об'єктів; можна легко поєднати із функціями втрат на основі регіону; може використовуватися в наборі даних із дисбалансом класів. Недоліки: складність обчислень; додаткове налаштування гіперпараметрів; нестабільна на зображеннях що мають шум, де важко обчислити контур.

Одні функції втрат добре себе показують для одних завдань, але погані в інших і навпаки. Для вирішення цього прийнято комбінувати функції помилок, або застосовувати принципи однієї до іншої. Наприклад, якщо ми хочемо, щоб *dice loss* більше фокусувався на об'єктах, які важко знайти (*hard values*), то ми можемо застосувати принципи *focal loss* тощо.

Отже, були розглянуті функції втрат: на основі розподілу, які порівнюють регіони об'єктів, які порівнюють межі об'єктів та комбіновані. Вибір функції втрат залежить від специфіки розв'язуваного завдання. Одні краще справляються з незбалансованим набором даних, другі швидші в обчисленнях, треті повільніше, але можуть точно обчислювати межі об'єкта, який треба знайти. Перераховані вище функції втрат підходять як для задач з одним класом, так і для задач з декількома класами, але є недолік, який полягає в тому, що вихідні результати для задач з декількома класами для підрахунку загальної помилки усереднюються. У такому разі може втрачатися частина інформації про розташування об'єкта, його важливість у конкретній сцені. Проблему втрати важливості об'єкта можна розв'язати введенням вагових коефіцієнтів, але так само потребує подальших досліджень проблема врахування взаємозв'язків об'єктів на сцені, оскільки перераховані функції цього не враховують. Ця проблема актуальна для розв'язання задачі сегментації дій та їхньої класифікації. Додатково варто дослідити функції втрат, які є одночасно і метриками. У такому випадку можна регулювати вибір даних для навчання, якщо знати відстань отриманого результату від еталону.

Список використаних джерел:

1. Bolya D. et al. Yolact: Real-time instance segmentation // Proceedings of the IEEE/CVF international conference on computer vision. – 2019. – С. 9157–9166.
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. Tversky A. Features of similarity // Psychological review. – 1977. – Т. 84. – №. 4. – С. 327.
4. Rezaeifighi H. et al. Generalized intersection over union: A metric and a loss for bounding box regression // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – С. 658–666.
5. Kervadec H. et al. Boundary loss for highly unbalanced segmentation // International conference on medical imaging with deep learning. – PMLR, 2019. – С. 285–296.