

ИСПОЛЬЗОВАНИЕ СИАМСКИХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ СХОЖЕСТИ ИЗОБРАЖЕНИЙ

© 2022 г. Е. А. ЗДОРОВЦОВА

Аннотация. В данной статье рассмотрено применение сиамской нейронной сети для решения задачи определения схожести изображений. Приведено сравнение модели с контрастной и триплетной функциями потерь.

Ключевые слова: сиамская нейросеть, siamese neural network, SNN, image similarity, triplet loss, contrastive loss, схожесть изображений.

ВВЕДЕНИЕ

В машинном обучении глубокие нейронные сети хорошо справляются почти со всеми задачами. Но для обучения и достижения хорошей точности они предполагают использование большого объема данных для каждого класса, чтобы показать хороший результат.

Однако для решения определенных задач (таких как распознавание лиц, проверка подлинности и др.) не всегда возможно рассчитывать на получение большего количества данных. Для решения такого рода задач есть другой тип архитектуры нейронной сети, который называется сиамская нейросеть. Сиамская нейронная сеть (СНР, Siamese Neural Network, SNN) — это разновидность нейронной сети, которая состоит из двух идентичных подсетей с одинаковыми наборами весов. Она является одним из наиболее популярных алгоритмов однократного обучения (one-shot learning). Это означает, что данная модель обычно используется, когда в каждом классе небольшое количество данных, но которых достаточно для получения точных прогнозов.

SNN не испытывают необходимости в большом наборе данных, даже при малом количестве выборок на каждый класс и несбалансированном распределении классов, они показывают достаточно высокую точность. Сиамские нейросети изучают функцию подобия. Таким образом, можно обучить модель решать задачу определения схожести изображений, т.е. отвечать на вопрос, совпадают ли два изображения. Это позволяет классифицировать новые классы данных без повторного обучения сети. В основном сиамская нейронная сеть имеет целью выяснить, насколько похожи две сопоставимые вещи, т.е. измерить вероятность того, что два входа принадлежат одному классу.

Как было показано ранее СНР состоит из двух идентичных подсетей. Каждая подсеть вычисляет характеристики одного входа, затем сходство признаков вычисляется с использованием функции потерь. Для входных пар одного и того же класса целевой выход равен 1, а для входных пар разных классов 0.

Мы обучаем сеть, чтобы минимизировать расстояние между выборками одного класса и увеличить расстояние между разными классами. Существует несколько видов функций подобия, с помощью которых можно обучать сиамскую сеть. В данной статье будут

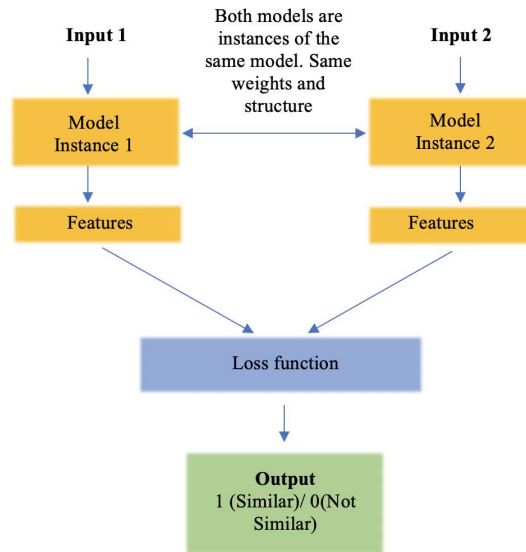


Рис. 1. Структура сиамской нейронной сети

рассмотрены две из них: контрастная потеря (contrastive loss) и потеря триплетов (triplet loss).

1. КОНТРАСТНАЯ ФУНКЦИЯ ПОТЕРЬ

Contrastive Loss – это функция потерь на основе расстояния. В данном случае рассматриваются пары изображений. Она позволяет различать являются два входных изображения похожими или нет. Контрастная потеря требует пары положительных и отрицательных обучающих данных. Целью данной функции потерь состоит в том, чтобы евклидово расстояние между представителями одного класса было меньше, чем между парой из разных классов. Функция потерь выглядит следующим образом:

$$L = Y * D^2 + (1 - Y) * \max \{margin - D, 0\}^2 \quad (1)$$

где D – евклидово расстояние, определяется как $D = \sqrt{((G(x_1) - G(x_2))^2)}$, $margin > 0$ – параметр, позволяющий отделять разные классы,

$$Y = \begin{cases} 1, & \text{если изображения } x_1, x_2 \text{ из одного класса} \\ 0, & \text{если изображения } x_1, x_2 \text{ из разных классов} \end{cases} \quad (2)$$

Нейросеть штрафует за отдаленность друг от друга изображений x_1, x_2 , если эти изображения на самом деле похожи. Аналогично, возникает штраф за близость изображений, которые не принадлежат одному классу.

2. ФУНКЦИЯ ПОТЕРЬ ТРИПЛЕТОВ

Triplet Loss – функция потерь, которая берет во внимание три объекта – якорь, положительный (похожий на якорь) и отрицательный объект (отличный от якоря). Это наиболее часто используемая функция потерь в задачах определения схожести. Здесь идет упор на минимизацию расстояния от якоря до позитива и максимизацию расстояния от якоря до негатива. Математически данную функцию потерь можно записать следующим образом:

$$L(A, P, N) = \max \{0, D(A, P) - D(A, N) + margin\}, \quad (3)$$

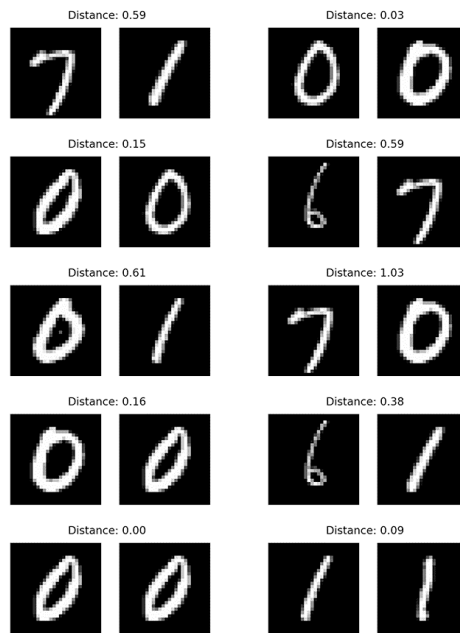


Рис. 2. Изображения с меньшими расстояниями предполагаются принадлежащими к одному классу, с большими – к разным

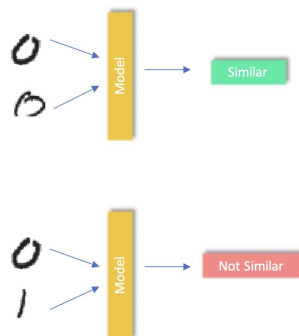


Рис. 3. Результат применения contrastive loss

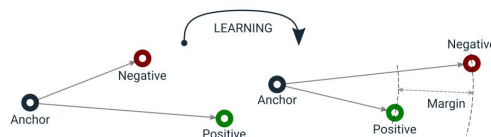


Рис. 4. Обучение функции потерь triplet loss для получение правильной кодировки

где A – якорь, P – позитив (изображение из того же класса, что и A , N – негатив (изображение из любого класса, отличного от класса якоря), $margin$ – параметр, позволяющий избежать сходимости к тривиальному решению. Т.к. модель может научиться создавать одинаковое кодирование для разных изображений (т.е. расстояния будут равны нулю), по этой причине добавляется гиперпараметр $margin$, чтобы всегда был разрыв между A и P по сравнению с A и N . Так триплетная функция потерь отвечает за то, чтобы разнородные пары были удалены от подобных пар, по крайней мере, на значение $margin$.

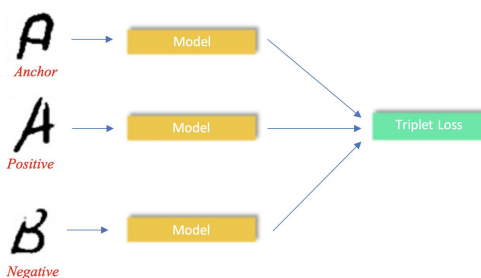


Рис. 5. Пример triplet loss

Основным преимуществом данной функции потерь над контрастной потерей состоит в том, triplet loss не имеет побочного эффекта, заключающегося в необходимости кодировать якорь и положительные образцы в одной и той же точке векторного пространства, как при применении контрастной потери. Это позволяет потере триплетов допускать некоторую внутрикласовую дисперсию, в отличие от контрастной потери, поскольку последняя сводит расстояние между якорем и любым положительным значением практически до 0.

Другими словами, потеря триплетов позволяет растягивать кластеры таким образом, чтобы включать выбросы, в то же время по-прежнему обеспечивая достаточный запас между образцами из разных кластеров.

ЗАКЛЮЧЕНИЕ

Сиамская сеть представляет собой класс нейросетей, состоящих из двух и более идентичных подсетей. Она позволяет определять сходство входных данных путем сравнения их векторов признаков, СНР основана на подходе к обучению метрик, который находит относительное расстояние между своими входами, используя контрастную потерю или триплетную потерю.

Сиамская нейросеть даже при малом количестве выборок на каждый класс показывает высокую точность, поэтому часто используется в задачах распознавания лиц, сравнения отпечатков пальцев и других задачах, где не кажется возможным собрать большое количество данных для каждого класса. Выходом сети является значение похожести входных изображений.

СПИСОК ЛИТЕРАТУРЫ

1. A friendly introduction to Siamese Networks [Электронный ресурс] / Sean Benhur J — Режим доступа: <https://towardsdatascience.com/a-friendly-introduction-to-siamese-networks-85ab17522942>.
2. One-Shot Learning With Siamese Network [Электронный ресурс] / Renu Khandelwal — Режим доступа: <https://medium.com/swlh/one-shot-learning-with-siamese-network-1c7404c35fda>.

Здоровцова Елена Александровна

Воронежский государственный университет, факультет прикладной математики и механики, кафедра программного обеспечения и администрирования информационных систем, магистр (ВГУ, ПММ, ПОиАИС)