

Контрастное обучение с изменчивыми внутриклассовыми эмбедингами

Китов В. В.^{1,2*}, Григорьев И. А.¹

^{1*} ф-т ВМК, МГУ имени М.В.Ломоносова, Ленинские горы, д.1,
стр. 52, 119991, Москва, Россия.

² дирекция по науке и инновациям, РЭУ имени Г.В.Плеханова,
Стремянный пер., д. 36, Москва, 109992, Москва, Россия.

* Автор(ы), ответственный(ые) за переписку. E-mail(s):

v.v.kitov@yandex.ru;

Соавторы: ilya.grigorev.00@inbox.ru;

Аннотация

В статье исследуется ограничение стандартных функций потерь для обучения сиамских нейронных сетей, заключающееся в требовании идентичности эмбедингов для объектов одного класса. Такая постановка задачи игнорирует естественную внутриклассовую вариативность данных и препятствует настройке нейросети полноценно решать целевую задачу классификации. В качестве решения предлагаются модификации попарной, тройной и вероятностной функции потерь, которые используются контрастным обучении. Ключевая идея состоит во введении гиперпараметра, задающего допустимый порог внутриклассового расстояния, ниже которого штраф не применяется. Этот подход позволяет нейронной сети сохранять информацию о различиях объектов одного класса и фокусироваться на основной задаче — разделении разных классов. Эффективность предложенных модификаций подтверждена экспериментально на различных наборах данных, где было продемонстрировано стабильное улучшение точности классификации. Также исследована зависимость точности прогнозов от нового гиперпараметра и от способа учёта эмбедингов при построении прогнозов.

Ключевые слова: классификация, нейронные сети, контрастное обучение, сиамские сети.

Получено редакцией 29.10.2025; внесены авторские правки 06.11.2025;
принята к публикации 07.11.2025

1 Введение

В работе рассматривается решение задачи классификации с помощью специальной сети-энкодера, называемой *сиамской сетью* (Siamese network), которая переводит объекты в их векторные представления (*эмбединги*) перед построением прогноза. Чтобы полученные эмбединги максимально подходили для задачи классификации, используется *контрастное обучение* (contrastive learning [1]), основанное на сопоставлении наборов объектов в пространстве признаков. Его идея заключается в том, чтобы *сближать* векторные представления объектов одного класса и *раздвигать* представления объектов разных классов. В результате подобного обучения объекты в пространстве эмбедингов станут группироваться по классам, а сама задача классификации сведётся к применению метрического классификатора в этом пространстве, используя один из подходов:

1. **Метод ближайших центроидов:** вычислить центроиды для каждого класса, усредняя эмбединги объектов соответствующего класса в обучающей выборке; тогда новый тестовый объект назначается тому классу, к центроиду которого он оказался ближе всего в пространстве эмбедингов [2].
2. **К ближайших соседей:** назначать объект тому классу, который встречается чаще всего среди K ближайших к нему объектов обучающей выборки в пространстве эмбедингов [3].

Контрастное обучение позволяет улучшить качество традиционных классификаторов прямого распространения [4] за счёт обучения не на отдельных объектах (как обычные классификаторы), а на *наборах объектов*, эмбединги которых сближаются/противопоставляются друг другу в зависимости от их классов. Поскольку наборов объектов больше, чем самих объектов, это приводит к более разнообразным обучающим примерам и более эффективному использованию обучающих данных. В частности, это позволяет сети обучаться отделять класс даже в случае, когда этот класс представлен всего несколькими обучающими примерами (*few-shot learning* [5]). Полученные с помощью сиамской сети эмбединги также дают возможность получения новых признаков представлений объектов, которые могут использоваться в родственных задачах (*transfer learning*).

Прогнозирование с помощью сиамских сетей активно используются в самых разных областях, например, для классификации изображений общего вида [6] и медицинских изображений [7], обработки потоковых данных [8], локализации роботов [9], оценки качества музыкальных исполнений [10] и распознавании человека по биометрическим данным [3].

Параметрами, настраиваемыми с помощью контрастного обучения являются веса сиамской сети, осуществляющей перевод объекта в соответствующий ему эмбединг так, что эмбединги объектов одного класса оказываются похожими, а эмбединги объектов разных классов — нет.

Для настройки сиамских сетей в основном используются *парные потери* (contrastive loss) [11], в которых сэмпляются пары объектов, принадлежащие либо одному, либо разным классам. Эти объекты пропускаются через сиамскую сеть, а затем потери штрафуют похожесть/расхождение эмбедингов, если объекты принадлежат одному/разным классам.

Другой популярной функцией потерь являются *тройные потери* (triples loss) [12], в которой сэмпляются уже тройки объектов: опорный объект, положительный объект (из того же класса, что и опорный) и отрицательный объект (принадлежащий другому классу). Эти объекты так же пропускаются через сиамскую сеть, а затем потери штрафуют разницу между расстояниями от опорного до положительного и от опорного до отрицательного объекта, добиваясь того, чтобы второе расстояние было заведомо больше первого.

Еще одна популярная функция потерь для настройки сиамских сетей — это *вероятностная функция потерь* (InfoNCE loss) [13, 14], в которой сэмпляются опорный объект, положительный и *набор* отрицательных объектов. Функция потерь поощряет сходство опорного и положительного объекта на контрасте с отрицательными объектами, которые должны оказываться непохожими на опорный объект.

Для обучения сиамских нейронных сетей применяют и другие функции потерь, такие как hinge loss [15], angular loss [16], multi-similarity loss [17], но они не настолько популярны, поэтому не будут затрагиваться в данном исследовании.

В контексте задачи классификации все рассмотренные функции потерь поощряют два условия:

- объектам разных классов должны назначаться непохожие эмбединги;
- объекты одного класса должны оказываться неотличимыми в пространстве эмбедингов.

Ключевым наблюдением работы является то, что даже в рамках одного класса объекты будут оказываться непохожими, поэтому требование идентичности внутриклассовых эмбедингов является излишним. Например, на Рис. 1 проиллюстрированы различные цифры в задаче классификации рукописных цифр. Видно, что каждая цифра может быть записана по-разному. В другом примере классификации изображений на кошек и собак, показанном на Рис. 2, кошки и собаки отличаются друг от друга размерами, окрасом, породой. Аналогично в медицинской задаче классификации болезни по симптомам, симптомы могут быть по-разному представлены при постановке одинакового диагноза.

Учитывая вариабельность среди объектов *одного класса*, нет смысла требовать идеальной схожести и для их эмбедингов, что заложено в классические функции потерь для настройки сиамих нейронных сетей — это только сместит оптимизацию с решения действительно важной задачи разделения эмбедингов для объектов разных классов, которая в основном и определяет точность итоговой классификации.

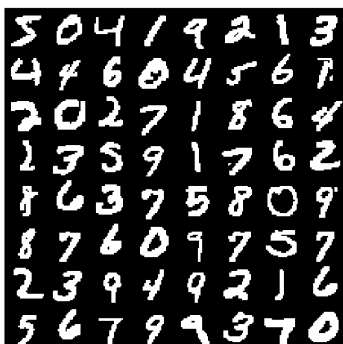


Рис. 1: Классификация рукописных цифр. Одинаковые цифры могут быть записаны по-разному.



Рис. 2: Классификация изображений на кошек и собак. Кошки, как и собаки, сильно различаются по внешнему виду.

Поэтому в работе предлагаются модификации попарных, тройных и вероятностных потерь таким образом, чтобы не штрафовались различия в эмбедингах ниже определённого уровня, определяемого вводимым гиперпараметром *естественной внутриклассовой вариабельности*. Предложенная модификация позволит переключить оптимизацию нейросети

со сближения внутриклассовых эмбеддингов в точки и больше сфокусироваться над разделением эмбеддингов для объектов разных классов, повысив точность классификации.

Другим вопросом, исследуемым в статье, является зависимость точности прогнозов от гиперпараметра внутриклассовой вариабельности, вводимого в модифицированных функциях потерь, а также выбор стратегии использования настроенных эмбеддингов: стоит ли использовать метод ближайших центроидов или метод K ближайших соседей?

Работа имеет следующую структуру. В главе 2 приводятся формулы для основных функций потерь, используемых в контрастном оценивании, и предлагаются их модификации. Глава 3 посвящена экспериментам по сравнению точности классификации для архитектур, настраиваемых стандартными и предложенными функциями потерь, а в главе 4 анализируются результаты проведённых экспериментов. Глава 5 посвящена изучению влияния гиперпараметра естественной внутриклассовой вариабельности на точность прогнозов, а в главе 6 приводятся заключительные выводы.

2 Обзор функций потерь и предлагаемые модификации

В данном разделе будут представлены классические функции потерь, применяемые для контрастного оценивания сямской нейронной сети $\mathbf{e} = f(\mathbf{x})$, преобразующей объекты в эмбеддинги так, чтобы объектам одного класса соответствовали похожие эмбеддинги, а эмбеддингам разных классов — отличающиеся.

Как будет видно из формул, эти потери поощряют максимальное сближение эмбеддингов объектов одного класса в полностью идентичные вектора. Учитывая внутриклассовую вариабельность объектов, предлагается каждую функцию потерь модифицировать таким образом, чтобы незначительные отличия в эмбеддингах объектов одного класса не подвергались штрафу, и приводятся формулы для модифицированных потерь. Степень допустимой внутриклассовой вариабельности будет контролироваться вводимым гиперпараметром естественной внутриклассовой вариабельности $v > 0$.

Попарные потери.

Попарные потери (contrastive loss [11]) вычисляются по парам сэмплируемых объектов $(\mathbf{x}_i, y_i); (\mathbf{x}_j, y_j)$ по следующей формуле:

$$L = (1 - Y) \cdot D^2 + Y \cdot \max\{0, m - D\}^2,$$

где

- $D = \rho(f(\mathbf{x}_i), f(\mathbf{x}_j))$ — расстояние между эмбеддингами;

- $m > 0$ — гиперпараметр, задающий минимальную нештрафуемую величину расстояния между эмбедингами объектов разных классов.
- $Y = \mathbb{I}\{y_i \neq y_j\}$, то есть $Y = 0$, если пара объектов из одного класса и $Y = 1$, если пара объектов из разных классов.

Из формулы данной функции потерь видно, что она стремится:

- раздвинуть эмбединги объектов разных классов на расстояние не ниже m ;
- свести расстояние между эмбедингами объектов одного класса к нулю.

Как было сказано выше, последнее свойство является нежелательным, поэтому предлагается следующая модификация функции потерь:

$$L = (1 - Y) \cdot \max\{0, D - v\}^2 + Y \cdot \max\{0, m - D\}^2$$

Модифицированная функция потерь уже не будет штрафовать различия эмбедингов объектов одного класса, если расстояние между ними меньше гиперпараметра естественной внутрикласовой вариации $v \geq 0$, причём при $v = 0$ модифицированные потери сводятся к исходным. Из смысла задачи $v < m$, иначе классы не будут разделяться.

Тройные потери.

Тройные потери (triplet loss [12]) считаются уже не по парам, а по тройкам сэмплируемых объектов:

1. опорный объект \mathbf{x} ;
2. положительный объект \mathbf{x}^+ из того же класса, что и опорный;
3. отрицательный объект \mathbf{x}^- из другого класса, нежели опорный.

Пусть

- $D_+ = \rho(f(\mathbf{x}), f(\mathbf{x}^+))$ — расстояние между эмбедингами опорного и положительного объекта;
- $D_- = \rho(f(\mathbf{x}), f(\mathbf{x}^-))$ — расстояние между эмбедингами опорного и отрицательного объекта.

Тогда тройная функция потерь вычисляется по формуле:

$$L = \max\{D_+^2 - D_-^2 + m, 0\},$$

где $m > 0$ — гиперпараметр, задающий насколько расстояния между эмбедингами объектов непохожих классов должны быть больше, чем расстояния между эмбедингами объектов одного класса.

Данная функция потерь поощряет ситуацию, когда межклассовые расстояния D_- оказываются большими, а внутрикласовые расстояния D_+ — малыми. Это может сдвигать решение не в сторону разделения классов на большие расстояния, а в сторону сближения внутрикласовых расстояний к нулю. Поскольку мы допускаем некоторый естественный

уровень внутриклассовой вариации эмбедингов $v > 0$, который не должен оказывать влияния на функцию потерь, то потери предлагается модифицировать следующим образом:

$$L = \max\{\max\{D_+, v\}^2 - D_-^2 + m, 0\}$$

В предлагаемой модификации внутриклассовым эмбедингам разрешается различаться на гиперпараметр внутриклассовой вариации $v \geq 0$, и это не оказывает влияния на минимизируемый критерий. При $v = 0$ модифицированные потери сводятся к исходным. Исходя из того, что эмбединги объекты разных классов должны быть разделены сильнее, чем эмбединги одного класса, следует, что $v < m$.

Вероятностные потери.

Вероятностная функция потерь InfoNCE (Noise Contrastive Estimation [13, 14]) вычисляется уже по набору из $R + 2$ объектов из обучающей выборки: опорный объект \mathbf{x} , положительный объект \mathbf{x}^+ (того же класса, что и опорный) и R отрицательных объектов $\mathbf{x}_1^-, \mathbf{x}_2^- \dots \mathbf{x}_R^-$ (отличающиеся по классу от опорного), а R — гиперпараметр.

Сама функция потерь считается по следующей формуле:

$$L = -\log \left(\frac{e^{s(f(\mathbf{x}^+), f(\mathbf{x}))}}{e^{s(f(\mathbf{x}^+), f(\mathbf{x}))} + \sum_{i=1}^R e^{s(f(\mathbf{x}_i^-), f(\mathbf{x}))}} \right),$$

где $s(\mathbf{e}, \mathbf{e}')$ - функция близости между эмбедингами, в качестве которой обычно берётся косинусная мера близости:

$$s(\mathbf{e}, \mathbf{e}') = \frac{\mathbf{e}^T \mathbf{e}'}{\|\mathbf{e}\| \cdot \|\mathbf{e}'\|}$$

Вероятностная функция потерь также реализует принцип контрастного обучения — она принимает наименьшие значения, когда числитель максимизирует сходство между внутриклассовыми эмбедингами, а знаменатель минимизирует сходство межклассовых эмбедингов.

Для того, чтобы сместить акцент настройки с минимизации внутриклассовых различий, которые в любом случае всегда будут присутствовать, вводится порог v , выше которого различия в похожести внутриклассовых эмбедингов уже не будут различаться критерием.

А саму функцию потерь предлагается модифицировать следующим образом:

$$L = -\log \left(\frac{e^{\min\{v, s(f(\mathbf{x}^+), f(\mathbf{x}))\}}}{e^{\min\{v, s(f(\mathbf{x}^+), f(\mathbf{x}))\}} + \sum_{i=1}^R e^{s(f(\mathbf{x}_i^-), f(\mathbf{x}))}} \right)$$

Из области значений косинусной меры близости имеет смысл рассматривать только значения $0 < v \leq 1$, а при $v = 1$ получаем исходную функцию потерь.

3 Сравнение точности методов

Для сравнения точности классификации классических потерь с предложенными модификациями были выбраны следующие датасеты, содержащие табличные данные и изображения:

- MNIST – 10-классовая классификация изображений рукописных цифр размера 28×28 пикселей представленный 10 классами [18].
- Fashion MNIST – 10-классовая классификация изображений одежды 28×28 пикселей [19].
- KMNIST – 10-классовая классификация изображений иероглифов 28×28 пикселей [20].
- Forest covertypes – 7-классовая классификация типа лесного покрова на основе картографических переменных [21].
- CIFAR-10 – классификация цветных изображений 32×32 пикселей на 10 классов [22].
- Heart Disease Dataset – бинарная классификация наличия/отсутствия сердечного заболевания у пациентов с заданными симптомами [23].
- Titanic Dataset – бинарная классификация выживания пассажиров с заданными характеристиками на корабле Titanic после катастрофы [24].

Для всех датасетов, кроме CIFAR-10, использовалась сиамская сеть, показанная на Рис. 3, на которой PReLU обозначает нелинейность parametric ReLU, а input обозначает размерность входного объекта.



Рис. 3: Схема сиамской сети для получения эмбедингов.

Категориальные признаки в табличных данных Forest covertypes, Heart Disease Dataset и Titanic Dataset кодировались с помощью one-hot кодирования. Для обучения на этих датасетах использовались мини-батчи размера 200, а размерность извлекаемых эмбедингов бралась равной 10.

Поскольку в датасете CIFAR-10 представлены цветные изображения большего размера, а сложность распознаваемых объектов выше, для него использовались эмбединги размера 20, а для извлечения эмбедингов использовалась предобученная сверточная нейронная сеть ResNet-18 [25] для работы с цветными изображениями 32×32 пикселя. Последний слой был заменен на полносвязный слой из 512 в 20 нейронов. Веса свёрточной сети так же дообучались в процессе настройки. Размер мини-батча для этого датасета был выбран 128.

Для настройки моделей использовался метод оптимизации Adam с шагом обучения 0.001, обучение длилось 5 эпох. Объекты в мини-батчах

выбирались случайным образом, но так, чтобы обеспечить совпадение/несовпадение классов в контрастных функциях потерь, в которых использовались следующие гиперпараметры:

- попарные и контрастные — $m = 2, v = 0.2$;
- вероятностные — $v = 0.5, K = 20$.

Для классификации по эмбедингам объектов использовался метод ближайшего центра. Для каждого класса мы храним его центр в эмбединговом пространстве, посчитанный как среднее по эмбедингам объектов обучающей выборки заданного класса. Далее для полученного эмбединга тестового объекта метод находит, к какому центру какого класса он расположен ближе всего. Класс найденного центра и является прогнозом. Сравнительные результаты точности классификации методом ближайшего центра на изображениях приведены в Таблице 1, а для табличных данных — в Таблице 2.

Таблица 1: Точность классификации изображений, используя метод ближайших центров.

| Функция потерь | MNIST | Fashion MNIST | KMNIST | CIFAR-10 |
|----------------------|---------------|---------------|---------------|---------------|
| попарная | 0.9757 | 0.8464 | 0.9043 | 0.8126 |
| модиф. попарная | 0.9765 | 0.8514 | 0.9135 | 0.8387 |
| тройная | 0.9492 | 0.7746 | 0.7334 | 0.7924 |
| модиф. тройная | 0.9561 | 0.7829 | 0.7642 | 0.8091 |
| вероятностная | 0.9520 | 0.8006 | 0.8573 | 0.8269 |
| модиф. вероятностная | 0.9637 | 0.8186 | 0.8769 | 0.8554 |

Таблица 2: Точность классификации табличных данных, используя метод ближайших центров.

| Функция потерь | Forest covertypes | Heart Disease | Titanic |
|----------------------|-------------------|---------------|---------------|
| попарная | 0.7638 | 0.7824 | 0.7401 |
| модиф. попарная | 0.7840 | 0.8091 | 0.7484 |
| тройная | 0.6384 | 0.7203 | 0.6572 |
| модиф. тройная | 0.6427 | 0.7472 | 0.6613 |
| вероятностная | 0.7142 | 0.7789 | 0.7532 |
| модиф. вероятностная | 0.7287 | 0.8025 | 0.7663 |

Дополнительно были проведены эксперименты классификацией на основе получаемых эмбедингов методом K ближайших соседей, при котором объект назначался тому классу, которому принадлежит K ближайших к нему обучающих объектов в эмбединговом пространстве.

В наших экспериментах использовалось $K = 5$. Результаты для классификации изображений приведены в Таблице 3, а для классификации табличных данных — в Таблице 4.

Таблица 3: Точность классификации изображений, используя метод 5 ближайших соседей.

| Функция потерь | MNIST | Fashion MNIST | KMNIST | CIFAR-10 |
|----------------------|---------------|---------------|---------------|---------------|
| попарная | 0.9781 | 0.8511 | 0.9078 | 0.8157 |
| модиф. попарная | 0.9795 | 0.8557 | 0.9174 | 0.8213 |
| тройная | 0.9558 | 0.7821 | 0.7420 | 0.7944 |
| модиф. тройная | 0.9604 | 0.7918 | 0.7701 | 0.8145 |
| вероятностная | 0.9553 | 0.8053 | 0.8623 | 0.8298 |
| модиф. вероятностная | 0.9671 | 0.8239 | 0.8809 | 0.8584 |

Таблица 4: Точность классификации табличных данных, используя метод 5 ближайших соседей.

| Функция потерь | Forest covertypes | Heart Disease | Titanic |
|----------------------|-------------------|---------------|---------------|
| попарная | 0.7672 | 0.7856 | 0.7432 |
| модиф. попарная | 0.7872 | 0.8125 | 0.7514 |
| тройная | 0.6448 | 0.7235 | 0.6613 |
| модиф. тройная | 0.6467 | 0.7502 | 0.6648 |
| вероятностная | 0.7175 | 0.7824 | 0.7563 |
| модиф. вероятностная | 0.7315 | 0.8057 | 0.7698 |

4 Анализ экспериментов

Из результатов экспериментов видно устойчивое преимущество в точности работы за счёт использования в контрастном обучении модифицированных потерь вместо базовых:

- модифицированные попарные потери работают лучше классических попарных;
- модифицированные тройные потери работают лучше классических тройных;
- модифицированные вероятностные потери работают лучше обычных вероятностных.

Модификации функций потерь позволяют сиамской сети больше переориентироваться на разделение объектов разных классов в эмбединговом пространстве вместо сжатия эмбедингов объектов одного класса в точки. В этом состоит основной результат работы.

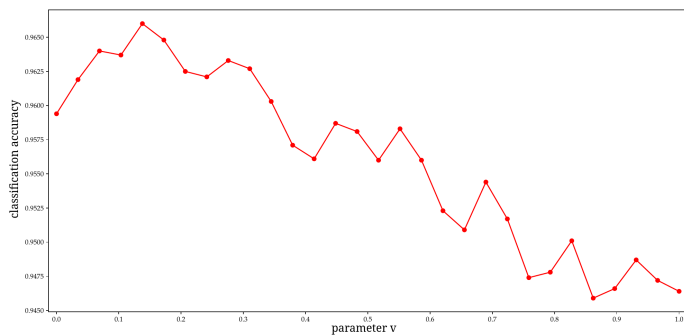


Рис. 4: Зависимость точности классификации с попарными потерями от v при $m = 2$ на датасете MNIST.

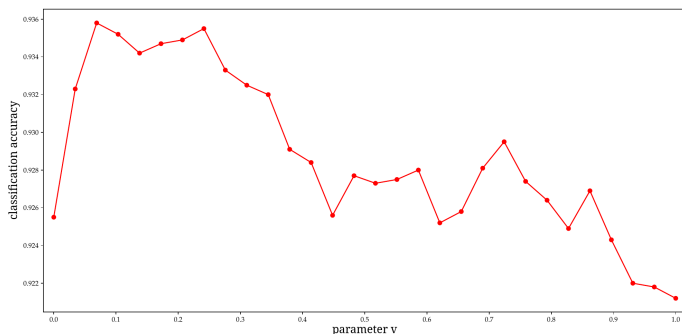


Рис. 5: Зависимость точности классификации с тройными потерями от v при $m = 2$ на датасете MNIST.

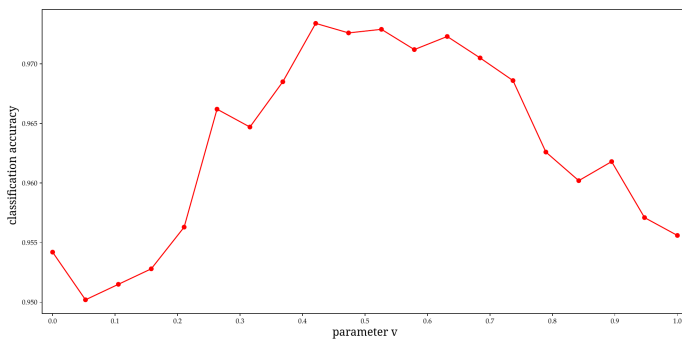


Рис. 6: Зависимость точности классификации с вероятностными потерями от v при $K = 20$ на датасете MNIST.

Также было проведено сравнение алгоритмов выбора метки класса на этапе применения модели. Метод К ближайших соседей лучше всего работал с $K = 5$ и, как видно из результатов, стабильно даёт более точные предсказания, чем метод ближайшего центроида. Мы это связываем с его возможностью извлекать нелинейные границы между классами в отличие от метода ближайшего центроида, в котором границы между любой парой классов оказываются линейными.

Среди всех видов функций потерь контрастного обучения наилучший результат достигался с попарными потерями кроме датасетов CIFAR-10 и Titanic, где более высокую точность обеспечило использование вероятностных потерь.

5 Влияние уровня внутриклассовой вариабельности

Изучим, как введённый гиперпараметр внутриклассовой вариабельности v влияет на точность прогнозов на примере датасета MNIST. В качестве итогового классификатора по эмбедингам будем использовать метод ближайших центроидов. Гиперпараметр m возьмём равным 2.

Зависимость точности классификации при настройке сиамской сети с помощью попарной, тройной и вероятностной функции потерь показана на Рис. 4, 5 и 6 соответственно.

Модифицированные попарные и тройные потери сводятся к исходным при $v = 0$. По графикам 4 и 5 видно, что оптимальная точность достигается при *положительных* значениях v , что обосновывает целесообразность модификации. Также по графикам видно, что при $v > 0.3$ точность начинает снижаться, поскольку эмбединги объектов разных классов становятся хуже разделимыми друг от друга.

Модифицированные вероятностные потери с косинусной мерой близости сводятся к исходным при $v = 1$. По графику 6 видно, что оптимальное значение $v \sim 0.5 < 1$, что также обосновывает целесообразность модификации.

6 Заключение

В работе рассмотрено решение задачи классификации с помощью сиамских нейронных сетей, настраиваемых методами контрастного обучения. Предложены модификации стандартных попарных, тройных и вероятностных потерь для этой задачи и проведены сравнительные эксперименты, которые показали, что модифицированные функции потерь устойчиво повышают точность классификации, а классифицировать по эмбедингам объектов лучше, используя метод К ближайших соседей, а не метод ближайших центроидов.

Авторские декларации

Финансирование

Работа выполнена в рамках государственного задания в сфере научной деятельности Министерства науки и высшего образования РФ на тему «Модели, методы и алгоритмы искусственного интеллекта в задачах экономики для анализа и стилизации многомерных данных, прогнозирования временных рядов и проектирования рекомендательных систем», номер проекта FSSW-2023-0004.

Доступность данных и программного кода

Код экспериментов доступен по ссылке ¹.

Конфликт интересов

Отсутствует.

Вклад авторов

В. В. Китов – постановка задачи и основная идея, написание текста статьи.

И. А. Григорьев – формализация функций потерь, программная реализация и проведение экспериментов, первоначальное оформление исследования.

Список литературы

- [1] *Bromley J. et al.* Signature verification using a "siamese" time delay neural network // *Advances in Neural Information Processing Systems*. — Vol. 6. — 1993.
- [2] *Lee D. Y., Lee M., Lee S. W.* Classification of imagined speech using Siamese neural network // *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. — IEEE, 2020. — Pp. 2979–2984.
- [3] *Hossain Progga P. et al.* A Bidirectional Siamese Recurrent Neural Network for Accurate Gait Recognition Using Body Landmarks // *arXiv e-prints*. — 2024.
- [4] *Khosla P. et al.* Supervised contrastive learning // *Advances in Neural Information Processing Systems*. — Vol. 33. — 2020. — Pp. 18661–18673.

¹<https://github.com/iluha-grig/MSU-CMC-master-thesis/blob/main/SiameseNN.ipynb>

- [5] *Zhu J. et al.* A few-shot meta-learning based siamese neural network using entropy features for ransomware classification // *Computers & Security*. — 2022. — Vol. 117. — P. 102691.
- [6] *Chen T. et al.* A simple framework for contrastive learning of visual representations // *International Conference on Machine Learning*. — PMLR, 2020. — Pp. 1597–1607.
- [7] *Battle M. L., Atapour-Abarghouei A., McGough A. S.* Siamese neural networks for skin cancer classification and new class detection using clinical and dermoscopic image datasets // *Proceedings of the 2022 IEEE International Conference on Big Data (Big Data)*. — IEEE, 2022. — Pp. 4346–4355.
- [8] *Malialis K., Panayiotou C. G., Polycarpou M. M.* Data-efficient online classification with siamese networks and active learning // *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*. — IEEE, 2020. — Pp. 1–7.
- [9] *Cabrera J. J. et al.* An experimental evaluation of Siamese Neural Networks for robot localization using omnidirectional imaging in indoor environments // *Artificial Intelligence Review*. — 2024. — Vol. 57, no. 8. — P. 198.
- [10] *Li X. et al.* Siamese residual neural network for musical shape evaluation in piano performance assessment // *Proceedings of the 2023 31st European Signal Processing Conference (EUSIPCO)*. — IEEE, 2023. — Pp. 1–5.
- [11] *Hadsell R., Chopra S., LeCun Y.* Dimensionality reduction by learning an invariant mapping // *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. — Vol. 2. — IEEE, 2006. — Pp. 1735–1742.
- [12] *Schroff F., Kalenichenko D., Philbin J.* Facenet: A unified embedding for face recognition and clustering // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. — 2015. — Pp. 815–823.
- [13] *Gutmann M., Hyvarinen A.* Noise-contrastive estimation: A new estimation principle for unnormalized statistical models // *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. — JMLR Workshop and Conference Proceedings, 2010. — Pp. 297–304.

- [14] *Chen T. et al.* A simple framework for contrastive learning of visual representations // *International Conference on Machine Learning.* — 2020. — Pp. 1597–1607.
- [15] *Weinberger K. Q., Saul L. K.* Distance metric learning for large margin nearest neighbor classification // *Journal of Machine Learning Research.* — 2009. — Vol. 10, no. 2.
- [16] *Wang J. et al.* Deep metric learning with angular loss // *Proceedings of the IEEE International Conference on Computer Vision.* — 2017. — Pp. 2593–2601.
- [17] *Wang X. et al.* Multi-similarity loss with general pair weighting for deep metric learning // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* — 2019. — Pp. 5022–5030.
- [18] MNIST dataset. — <https://www.kaggle.com/datasets/hojjatk/mnist-dataset>. — (дата обращения: 26.08.2025).
- [19] *Xiao H., Rasul K., Vollgraf R.* Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms // *arXiv preprint arXiv:1708.07747.* — 2017.
- [20] KMNIST dataset. — <https://datasets.activeloop.ai/docs/ml/datasets/kmnist/>. — (дата обращения: 26.08.2025).
- [21] Forest cover types dataset. — <https://www.kaggle.com/datasets/uciml/forest-cover-type-dataset/data>. — (дата обращения: 26.08.2025).
- [22] CIFAR-10 dataset. — <https://www.kaggle.com/datasets/ayush1220/cifar10>. — (дата обращения: 26.08.2025).
- [23] Heart Disease Dataset. — <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>. — (дата обращения: 26.08.2025).
- [24] Titanic Dataset. — <https://www.kaggle.com/datasets/yasserh/titanic-dataset>. — (дата обращения: 26.08.2025).
- [25] *He K. et al.* Deep residual learning for image recognition // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* — 2016. — Pp. 770–778.