

УДК 004.89
DOI: 10.15827/0236-235X.137.028-036

Дата подачи статьи: 30.09.21, после доработки: 18.10.21
2022. Т. 35. № 1. С. 028–036

Метод адаптивной классификации изображений с использованием обучения с подкреплением

А.А. Елизаров¹, магистрант, artelizar@gmail.com

¹ Казанский (Приволжский) федеральный университет, г. Казань, 420008, Россия

В статье представлен метод классификации изображений с использованием, помимо базовой нейронной сети, дополнительной, способной адаптивно концентрироваться на классифицируемом объекте изображения. Задача дополнительной сети является задачей о контекстном многооруком бандите и сводится к предсказанию такой области на исходном изображении, при вырезании которой в процессе классификации возрастет уверенность базовой нейронной сети в принадлежности объекта на изображении правильному классу. Обучение дополнительной сети происходит с помощью методов обучения с подкреплением и стратегий достижения компромисса между эксплуатацией и исследованием при выборе действий для решения задачи о контекстном многооруком бандите.

На подмножестве набора данных ImageNet-1K проведены различные эксперименты по выбору архитектуры нейронной сети, алгоритма обучения с подкреплением и стратегии исследования при обучении. Рассмотрены такие алгоритмы обучения с подкреплением, как DQN, REINFORCE и A2C, и такие стратегии исследования, как ϵ -жадная, ϵ -softmax, ϵ -decay-softmax и метод UCB1. Большое внимание уделено описанию проведенных экспериментов и обоснованию полученных результатов.

Предложены варианты применения разработанного метода, демонстрирующие увеличение точности классификации изображений по сравнению с базовой моделью ResNet. Дополнительно рассмотрен вопрос о вычислительной сложности данного метода.

Дальнейшие исследования могут быть направлены на обучение агента на изображениях, не задействованных при обучении сети ResNet.

Ключевые слова: искусственный интеллект, машинное обучение, нейронные сети, компьютерное зрение, классификация изображений, обучение с подкреплением, задача о контекстном многооруком бандите.

Многие современные интеллектуальные системы управления так или иначе оснащены модулями, которые получают информацию об окружающем мире посредством анализа изображений. В этом аспекте на первый план выходят задачи компьютерного зрения (прежде всего – классификации изображений) и методы их решения.

Задача классификации изображений заключается в разделении объектов на изображениях на группы, называемые классами [1]. Со времени проведения соревнования ILSVRC 2012 (ImageNet Large Scale Visual Recognition Challenge) [2] и по сегодня лучшие результаты в решении этой задачи показывают сверточные нейронные сети, превзошедшие в 2015 году уровень классификации изображений человеком на наборе данных ImageNet [3].

В последнее время для решения самых различных задач успешно применяется такое направление машинного обучения, как обучение с подкреплением. В обучении с подкреплением некоторая система (агент) обучается, взаимодействуя с окружающей средой и получая

от нее отклик [4], то есть обучение агента осуществляется методом проб и ошибок. Хотя основная теория современного обучения с подкреплением была разработана еще в прошлом веке, значительных успехов в этом направлении удалось добиться относительно недавно [5, 6]. В первую очередь это произошло благодаря развитию вычислительной техники и методов глубокого обучения. В силу того, что агент, как правило, не обладает полной информацией об окружающей среде, методы обучения с подкреплением активно применяются при разработке различных интеллектуальных информационных систем.

Одними из первых основных методов, использующих обучение с подкреплением при решении задачи классификации изображений и показывающих в настоящее время результаты на уровне state-of-the-art, являются FCAN [7] и DT-RAM [8]. В методе FCAN нейронная сеть имитирует зрительную систему человека, которая распознает объект через серию проблесков на наиболее важных частях этого объекта, отличающих его от других. На каждом времен-

ном шаге нейронная сеть выбирает определенную область на входном изображении и с учетом областей, выбранных на предыдущих шагах, пытается классифицировать объект на изображении. Для обучения такой сети используются метод обучения с подкреплением REINFORCE [4] и модифицированная жадная стратегия выбора действий. Идея метода DT-RAM основана на способности человека к избирательному зрительному вниманию, когда зрительные сигналы подавляют ненужные области, присутствующие в поле зрения человека, тем самым помогая ему в поиске цели. В DT-RAM используется глубокая рекуррентная нейронная сеть с итеративным механизмом визуального внимания, для обучения которой применяется модифицированный алгоритм REINFORCE. К преимуществам методов FCAN и DT-RAM можно отнести использование при обучении только меток правильных классов без более подробной аннотации изображений.

К недавним работам по данной теме относятся, например, [9], в которой нейронная сеть обучается динамически определять необходимые для правильной классификации участки изображений с высоким разрешением на основе парных изображений с низким разрешением. Такой подход позволяет снизить вычислительную сложность метода и повысить точность классификации изображений. В [10] исследуется проблема классификации изображений с использованием нескольких агентов, способных самостоятельно собирать информацию с ограниченного участка изображения и обмениваться ею друг с другом. Экспериментальные результаты демонстрируют эффективность структуры мультиагентного обучения.

Целью данной работы является разработка метода классификации изображений с использованием, помимо базовой нейронной сети ResNet, дополнительной (агента), которая обучается с помощью методов обучения с подкреплением [11] и способна адаптивно концентрироваться на классифицируемом объекте изображения.

Описание метода адаптивной классификации изображений

Идея разработанного метода заключается в использовании, помимо нейронной сети для классификации изображений, дополнительной нейронной сети – агента. В качестве базовой модели для классификации изображений при-

меняется нейронная сеть ResNet-50 [12], предобученная на выбранном наборе данных. Агент получает на вход выход 3-й группы остаточных блоков ResNet и предсказывает, какую область из полученных карт признаков изображения нужно вырезать, чтобы в процессе классификации возросла уверенность сети ResNet в принадлежности объекта на исходном изображении правильному классу. После этого карты признаков изображения обрезаются согласно предсказанию агента и подаются на вход следующей группе остаточных блоков ResNet для последующего извлечения признаков и классификации. Для вырезания карт признаков изображения и приведения их к одному размеру используется слой ROI Pool [13], часто применяемый на практике при решении задач локализации объектов на изображениях.

Такая задача для агента ставится исходя из предположения, что при правильном вырезании сеть ResNet сможет лучше классифицировать объект на изображении, и является задачей о контекстном многоруком бандите [4], так как эпизод состоит только из одного шага и во время обучения агент получает m различных контекстов (m – количество изображений в обучающей выборке).

Состояниями окружающей среды являются карты признаков изображений – выходы 3-й группы остаточных блоков нейронной сети ResNet.

Действия агента – варианты вырезания части карт признаков изображения. Они формируются следующим образом: входные карты признаков изображения замощаются с дискретным шагом λ слева направо и сверху вниз каждым из ограничивающих прямоугольников с размерами сторон, равными соответственно 0.5, 0.75 и 1.0 от размера сторон входных карт признаков изображения. Замощение ограничивающим прямоугольником с размером сторон 1.0 от размера сторон карт признаков изображения является действием агента «ничего не вырезать».

Вознаграждение агента на временном шаге t для i -го изображения из набора данных можно записать как

$$r_{it} = outputs_{it}(c_i) - baseline_outputs_{it}(c_i),$$

где $outputs_{it}$, $baseline_outputs_{it}$ – векторы уверенности нейронной сети ResNet в принадлежности объекта на i -м изображении различным классам, полученные соответственно с учетом вырезания агентом карт признаков изображения и без учета вырезания (базовая модель ResNet, эквивалентно действию «ничего не вы-

резать»), а c_i – индекс правильного класса. Таким образом, вознаграждение агента r_{it} – вещественное число на отрезке от -1 до 1 .

Нейронная сеть агента обучается с помощью алгоритмов обучения с подкреплением с применением стратегий достижения компромисса между эксплуатацией и исследованием при выборе действий для решения задачи о многоруком бандите. Подразумевается, что при обучении нейронной сети агента веса сети ResNet заморожены. Для обучения агента в разработанном методе необходимы только метки правильных классов изображений без более подробной аннотации. Схема архитектуры разработанного метода приведена на рисунке 1.

Методика проведения экспериментов

При проведении экспериментов для обучения нейронной сети ResNet использовались изображения как из валидационной выборки агента, так и из обучающей. Этот выбор основан на предположении, что для эффективной работы агента нужно обучать и тестировать его на выходах группы блоков ResNet из одного распределения. Другими словами, и обучать, и тестировать агента необходимо на изображениях, которые либо применялись при обучении сети ResNet, либо нет. Для упрощения исследования влияния параметров на эффективность

работы агента сначала были проведены эксперименты с агентом, который обучался и тестировался на изображениях, использованных при обучении ResNet.

Описание набора данных

В настоящей работе для проведения экспериментов использовалось подмножество набора данных ImageNet-1K [14], содержащего почти 1.4 млн изображений с высоким разрешением 1 000 различных классов. В подмножество были включены только классы животных, чтобы приблизить исходную задачу классификации изображений к задаче детального распознавания (*fine-grained recognition*) [15]. Таким образом, может потенциально увеличиться влияние вырезания карт признаков изображений на эффективность классификации.

Основные характеристики использованного подмножества: 60 классов различных животных, 14 тыс. изображений для обучения нейронной сети ResNet, 55 тыс. изображений для обучения агента, 8 тыс. изображений в валидационной выборке.

Архитектура нейронной сети

В качестве архитектуры нейронной сети агента рассматривались варианты, представленные в таблицах 1–4. Архитектура из таб-

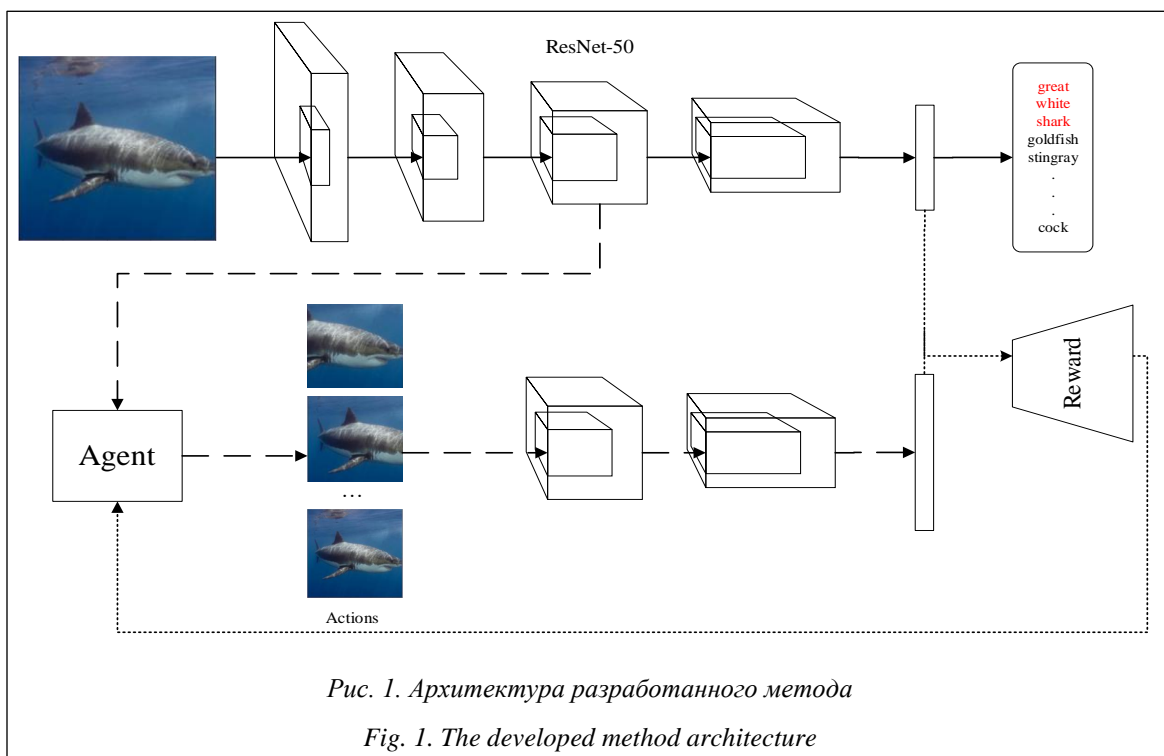


Рис. 1. Архитектура разработанного метода

Fig. 1. The developed method architecture

лицы 1 является базовым прототипом, а в архитектурах из таблиц 2–4 для уменьшения размерности входа используются слой субдискретизации MaxPool и сверточный слой с размером ядра свертки 1×1 и слоем нормализации BatchNorm. По умолчанию размер скрытого полносвязного слоя в этих архитектурах – 512 нейронов, а шаг замощения λ равен 2,

что приводит к 26 возможным действиям агента.

На рисунках 2 и 3 приведены графики зависимости точности классификации моделей на валидационной выборке от используемой архитектуры. Черной пунктирной линией обозначена точность классификации базовой модели ResNet-50 на валидационной выборке.

Таблица 1

Полносвязная архитектура

Table 1

The fully connected architecture

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14×14×1024
1	Полносвязный	ReLU	512	512
2	Полносвязный	–	26	26

Таблица 2

Архитектура с уменьшением размера карт признаков изображения до 7×7 или 3×3

Table 2

The architecture with image feature maps reduction to 7×7 or 3×3

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14×14×1024
1	MaxPool	–	–	7×7×1024 / 3×3×1024
2	Полносвязный	ReLU	512	512
3	Полносвязный	–	26	26

Таблица 3

Архитектура с уменьшением числа каналов карт признаков изображения до 512 или 256

Table 3

The architecture with reduction the number of channels of image feature maps to 512 or 256

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14×14×1024
1	Сверточный	–	1×1, 512 / 1×1, 256	14×14×512 / 14×14×256
2	BatchNorm	ReLU	512 / 256	14×14×512 / 14×14×256
3	Полносвязный	ReLU	512	512
4	Полносвязный	–	26	26

Таблица 4

Архитектура с уменьшением размера до 7×7 и числа каналов карт признаков изображения до 512

Table 4

The architecture with reduction of image feature maps to 7×7 and the number of channels of image feature maps to 512

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14×14×1024
1	MaxPool	–	–	7×7×1024
2	Сверточный	–	1×1, 512	7×7×512
3	BatchNorm	ReLU	512	7×7×512
4	Полносвязный	ReLU	512	512
5	Полносвязный	–	26	26

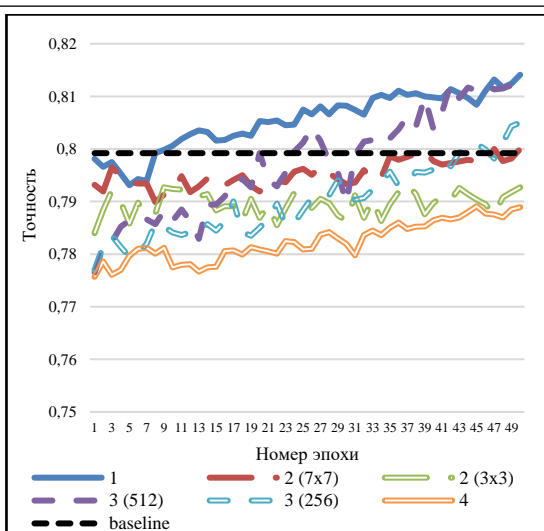


Рис. 2. Точность классификации моделей в зависимости от используемой архитектуры

Fig. 2. The model classification accuracy depending on the used architecture

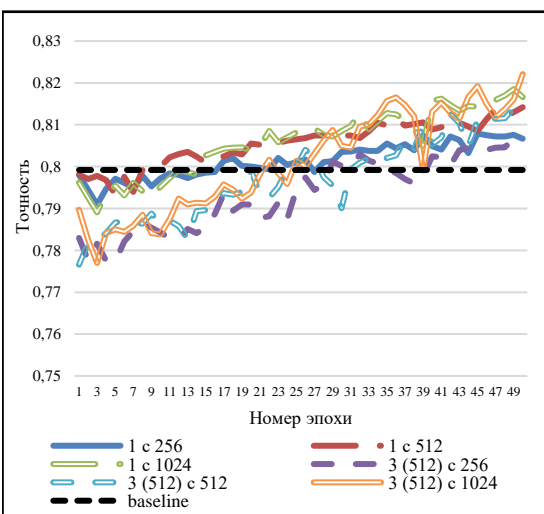


Рис. 3. Точность классификации лучших моделей в зависимости от числа нейронов в скрытом полносвязном слое

Fig. 3. The classification accuracy of the best models depending on the number of neurons in a hidden fully connected layer

Среди разработанных архитектур лучшие результаты показали полносвязная архитектура из таблицы 1 и архитектура с уменьшением числа каналов карт признаков изображения до 512 из таблицы 3. Среди этих моделей наилучший результат продемонстрировала архитектура из таблицы 3 с уменьшением числа каналов карт признаков изображения до 512 с 1 024 нейронами в скрытом полносвязном

слое. Она и была выбрана в качестве основной архитектуры для нейронной сети агента.

Следует отметить, что выбранная архитектура является неглубокой и содержит примерно 100 млн обучаемых параметров, что составляет около 80 % от количества обучаемых параметров основной нейронной сети ResNet-50. Поэтому применение агента незначительно увеличивает вычислительную сложность разработанной модели даже несмотря на то, что основная сеть ResNet сама по себе имеет невысокую вычислительную сложность. Более того, для увеличения итоговой точности классификации в качестве основной нейронной сети вместо ResNet-50 могут использоваться самые современные сверточные нейронные сети, являющиеся модификациями сетей ResNet, например, EfficientNet [16], которая на сегодняшний день показывает лучшие результаты в задаче классификации изображений на различных наборах данных.

Алгоритм обучения и стратегия выбора действий

В качестве алгоритмов для обучения агента рассматривались алгоритмы обучения с подкреплением DQN [6], REINFORCE и A2C [4]. Алгоритм DQN использовался с модификацией воспроизведения опыта [6] с максимально возможным размером буфера. В функциях потерь алгоритмов REINFORCE и A2C также применялся энтропийный бонус [17].

В качестве стратегий выбора действий (стратегий исследования) при обучении агента рассматривались следующие стратегии достижения компромисса между эксплуатацией и исследованием: ϵ -жадная (ϵ -greedy) [4], ϵ -softmax (модификация softmax [4] по аналогии с ϵ -жадной стратегией), ϵ -decay-softmax (по аналогии с ϵ -decay-жадной стратегией [4]) и метод UCB1 [18]. Параметр ϵ в стратегии ϵ -decay-softmax уменьшался в процессе обучения от 1.0 до 0.5.

Для каждого из алгоритмов была выбрана наиболее оптимальная стратегия исследования. В результате этого эксперимента построены графики для алгоритмов DQN (см. <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/2.jpg>, <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/3.jpg>), REINFORCE (см. <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/4.jpg>, <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/5.jpg>), A2C (см. <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/5.jpg>).

www.swsys.ru/uploaded/image/2022-1/2022-1-dop/6.jpg, <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/7.jpg>).

Проведенные эксперименты показали следующее.

- Результаты применения стратегии ϵ -decay-softmax выглядят немного лучше по сравнению с ϵ -softmax. Однако в целом эффективность этих стратегий в данной задаче практически не отличается. Это говорит о том, что большая вероятность исследования в начале обучения не повышает точность классификации.

- Результаты работы метода UCB1 (см. <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/2.jpg>, <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/4.jpg>, <http://www.swsys.ru/uploaded/image/2022-1/2022-1-dop/6.jpg>), несмотря на отсутствие случайности, выглядят несколько шумно по сравнению с результатами других стратегий. Это следует из того, что метод UCB1 довольно часто выбирает недоисследованные действия.

- В общем, использование в этой задаче алгоритмов обучения с подкреплением, основанных на функциях ценности, более эффективно, чем алгоритмов, основанных на функции политики.

Среди лучших моделей для всех алгоритмов обучения с подкреплением самый хороший результат показала модель, обученная в течение 150 эпох с помощью алгоритма DQN и ϵ -жадной стратегии выбора действий при $\epsilon = 0.5$. Она превысила точность классификации базовой модели ResNet-50 почти на 0.06, достигнув точности классификации изображений на валидационной выборке, равной 0.857. Итоговые результаты отражены на рисунке 4 и в таблице 5.

Таблица 5

Результаты классификации лучших моделей по сравнению с базовой моделью ResNet-50

Table 5

The results of classifying the best models compared to the ResNet-50 base model

Модель	Точность классификации на валидационной выборке
DQN с ϵ -жадной стратегией при $\epsilon = 0.5$	0.8571
DQN с методом UCB1	0.8514
A2C со стратегией ϵ -softmax при $\epsilon = 0.5$	0.8397
REINFORCE с ϵ -жадной стратегией при $\epsilon = 0.5$	0.8392
ResNet-50	0.7992

Благодаря более длительному времени обучения также хорошо зарекомендовала себя модель, обученная с помощью алгоритма DQN и метода UCB1. Это объясняется тем, что методу UCB1 требуется выполнить каждое действие для каждого изображения из обучающей выборки хотя бы по несколько раз, чтобы получить достаточно точные оценки ценности действий.

Способ обучения агента на данных, не использованных при обучении базовой сети ResNet

Эксперименты по выбору архитектуры агента, алгоритма обучения с подкреплением и стратегии исследования при обучении проводились при условии, что изображения из валидационной выборки агента использовались при обучении нейронной сети ResNet. На практике такой подход неприменим, но можно предложить другой, в котором валидационная выборка используемого набора данных делится на две части: на одной проходит обучение агента, а на другой его тестирование. Таким образом, нейронная сеть ResNet не будет обучаться на изображениях из обучающей и валидационной выборки агента.

Также был проведен эксперимент, где в качестве валидационной выборки использовалась валидационная выборка всего набора данных ImageNet-1K, которая содержит 50 тыс. изображений 1 000 различных классов. Она

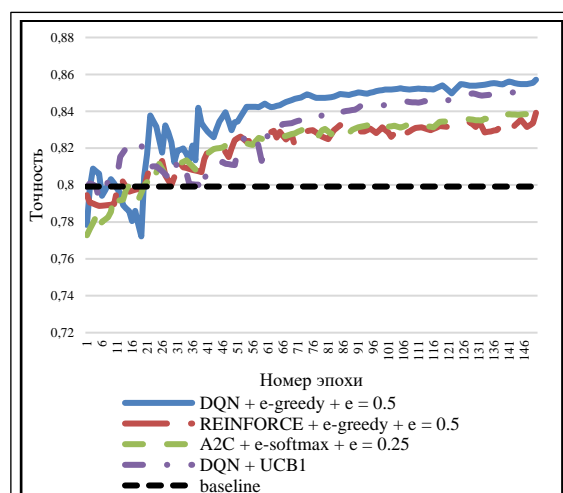


Рис. 4. Точность классификации лучших моделей по сравнению с базовой моделью ResNet-50

Fig. 4. The classification accuracy of the best models compared to the ResNet-50 base model

была поделена на две части: 45 тыс. изображений для обучения и 5 тыс. изображений для тестирования агента. В качестве параметров агента применялись полученные по результатам предыдущих экспериментов. По итогам этого эксперимента точность классификации на тестовой части валидационной выборки с учетом вырезания агентом карт признаков изображений составила 0.6942, что превышает точность базовой модели ResNet, равную 0.6934. Отметим, что обучение агента на изображениях, которые не использовались при обучении сети ResNet, может потребовать большего объема тренировочных данных, а также более детального подбора гиперпараметров.

Заключение

В статье описан метод классификации изображений с использованием дополнительной нейронной сети (агента), которая обучается с помощью методов обучения с подкреплением и способна адаптивно концентрироваться на классифицируемом объекте изображения. Для обучения разработанного метода необходимы только метки правильных классов изображений, что не усложняет применение его на практике по сравнению с базовой моделью ResNet.

Основная идея метода базируется на предположении, что вырезание агентом правильной области на картах признаков изображения в процессе классификации сетью ResNet может потенциально увеличить точность классификации. Такая идея приводит к задаче о контекстном многоруком бандите, где состояниями окружающей среды являются входные карты признаков изображений, действиями –

возможные области для вырезания на входных картах признаков, а вознаграждением – разность уверенностей сети ResNet в принадлежности объекта на изображении правильному классу с учетом вырезания агентом карт признаков изображения и если ничего не вырезать (базовая модель ResNet), соответственно.

Проведены эксперименты по выбору архитектуры агента, алгоритма обучения с подкреплением и стратегии исследования при обучении агента. Особое внимание уделено описанию проведенных экспериментов и обоснованию полученных результатов.

Эксперименты проводились при условии, что изображения из валидационной выборки агента используются при обучении сети ResNet, чтобы оценить эффективность работы агента для входа из выходов группы блоков ResNet из одного распределения. При таком подходе разработанная модель показала точность классификации, значительно превосходящую точность классификации базовой модели, что создает задел для будущих исследований. Однако этот подход неприменим на практике, поэтому был предложен другой, при котором изображения из обучающей и валидационной выборок агента не используются при обучении сети ResNet. Модифицированный подход также показал увеличение точности классификации изображений по сравнению с базовой моделью ResNet-50. Дальнейшие исследования могут быть связаны с обучением агента на изображениях, которые не были задействованы при обучении сети ResNet, так как это может потребовать большего объема тренировочных данных и более детального подбора гиперпараметров.

Автор выражает благодарность научному руководителю, к.ф.-м.н. Разинкову Е.В.

Литература

1. Goodfellow I., Bengio Y., Courville A. Deep learning. Genet Program Evolvable Mach, 2018, vol. 19, pp. 305–307. DOI: 10.1007/s10710-017-9314-z.
2. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017, vol. 60, no. 6, pp. 84–90. DOI: 10.1145/3065386.
3. Russakovsky O., Deng J., Su H. et al. ImageNet large scale visual recognition challenge. Int. J. of Computer Vision, 2015, vol. 115, no. 3, pp. 211–252. DOI: 10.1007/s11263-015-0816-y.
4. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. Cambridge, The MIT Press, 2014, 338 p.
5. Silver D., Huang A., Maddison C.J. et al. Mastering the game of Go with deep neural networks and tree search. Nature, 2016, vol. 529, no. 7587, pp. 484–489. DOI: 10.1038/nature16961.
6. Mnih V., Kavukcuoglu K., Silver D. et al. Playing Atari with deep reinforcement learning. ArXiv, 2013, art. 1312.5602. URL: <https://arxiv.org/abs/1312.5602> (дата обращения: 15.06.2021).
7. Liu X., Xia T., Wang J. et al. Fully convolutional attention networks for fine-grained recognition. ArXiv, 2017, art. 1603.06765. URL: <https://arxiv.org/abs/1603.06765> (дата обращения: 15.06.2021).

8. Li Z., Yang Y., Liu X. et al. Dynamic computational time for visual attention. Proc. IEEE ICCVW, 2017, pp. 1199–1209. DOI: 10.1109/ICCVW.2017.145.
9. Uzket B., Ermon S. Learning when and where to zoom with deep reinforcement learning. Proc. IEEE/CVF CVPR, 2020, pp. 12342–12351. DOI: 10.1109/CVPR42600.2020.01236.
10. Mousavi H.K., Nazari M., Takáč M. et al. Multi-Agent image classification via reinforcement learning. Proc. IEEE/RSJ IROS, 2019, pp. 5020–5027. DOI: 10.1109/IROS40897.2019.8968129.
11. Елизаров А.А., Разинков Е.В. Классификация изображений с использованием обучения с подкреплением // Электронные библиотеки. 2020. Т. 23. № 6. С. 1172–1191. DOI: 10.26907/1562-5419-2020-23-6-1172-1191.
12. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. Proc. IEEE/CVF CVPR, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.
13. Girshick R. Fast R-CNN. Proc. IEEE ICCV, 2015, pp. 1440–1448. DOI: 10.1109/ICCV.2015.169.
14. ImageNet Dataset. URL: <http://image-net.org> (дата обращения: 15.06.2021).
15. Papers with Code: Fine-Grained Image Classification. URL: <https://paperswithcode.com/task/fine-grained-image-classification> (дата обращения: 20.06.2021).
16. Mingxing T., Quoc V.L. EfficientNet: Rethinking model scaling for convolutional neural networks. ArXiv, 2019, art. 1905.11946v5. URL: <https://arxiv.org/pdf/1905.11946.pdf> (дата обращения: 15.06.2021).
17. Abdolmaleki A., Springenberg J.T., Degraeve J. et al. Relative entropy regularized policy iteration. ArXiv, 2018, art. 1812.02256. URL: <https://arxiv.org/abs/1812.02256> (дата обращения: 15.06.2021).
18. Auer P., Cesa-Bianchi N., Fischer P. Finite-time analysis of the multiarmed bandit problem. Machine Learning, 2002, vol. 47, pp. 235–256. DOI: 10.1023/A:1013689704352.

Software & Systems
DOI: 10.15827/0236-235X.137.028-036

Received 30.09.21, Revised 18.10.21
2022, vol. 35, no. 1, pp. 028–036

The adaptive image classification method using reinforcement learning

A.A. Elizarov¹, Graduate Student, artelizar@gmail.com

¹ Kazan (Volga Region) Federal University, Kazan, 420008, Russian Federation

Abstract. The paper proposes a method for image classification that uses in addition to a basic neural network for image classification an additional neural network able to adaptively concentrate on the classified image object. The task of the additional network is the contextual multi-armed bandit problem, which reduces to predicting such area on the original image, which is, when cut out of the classification process, will increase the confidence of the basic neural network that the object on the image belongs to the correct class. The additional network is trained using reinforcement learning techniques and strategies for compromising between exploration and research when choosing actions to solve the contextual multi-armed bandit problem.

Various experiments were carried out on a subset of the ImageNet-1K dataset to choose a neural network architecture, a reinforcement learning algorithm and a learning exploration strategy. We considered reinforcement learning algorithms such as DQN, REINFORCE and A2C and learning exploration strategies such as ϵ -greedy, ϵ -softmax, ϵ -decay-softmax and UCB1 method. Much attention was paid to the description of the experiments performed and the substantiation of the results obtained.

The paper proposes application variants of the developed method, which demonstrate an increase in the accuracy of image classification in comparison with the basic ResNet model. It additionally considers the issue of the computational complexity of the developed method.

Keywords: artificial intelligence, machine learning, neural networks, computer vision, image classification, reinforcement learning, contextual multi-armed bandit problem.

Acknowledgements. I would like to express my gratitude to my scientific advisor Ph.D. (Physics and Mathematics) E.V. Razinkov

References

1. Goodfellow I., Bengio Y., Courville A. Deep learning. *Genet Program Evolvable Mach*, 2018, vol. 19, pp. 305–307. DOI: 10.1007/s10710-017-9314-z.
2. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, vol. 60, no. 6, pp. 84–90. DOI: 10.1145/3065386.
3. Russakovsky O., Deng J., Su H. et al. ImageNet large scale visual recognition challenge. *Int. J. of Computer Vision*, 2015, vol. 115, no. 3, pp. 211–252. DOI: 10.1007/s11263-015-0816-y.
4. Sutton R.S., Barto A.G. *Reinforcement Learning: An Introduction*. Cambridge, The MIT Press, 2014, 338 p.
5. Silver D., Huang A., Maddison C.J. et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, vol. 529, no. 7587, pp. 484–489. DOI: 10.1038/nature16961.
6. Mnih V., Kavukcuoglu K., Silver D. et al. Playing Atari with deep reinforcement learning. *ArXiv*, 2013, art. 1312.5602. Available at: <https://arxiv.org/abs/1312.5602> (accessed June 15, 2021).
7. Liu X., Xia T., Wang J. et al. Fully convolutional attention networks for fine-grained recognition. *ArXiv*, 2017, art. 1603.06765. Available at: <https://arxiv.org/abs/1603.06765> (accessed June 15, 2021).
8. Li Z., Yang Y., Liu X. et al. Dynamic computational time for visual attention. *Proc. IEEE ICCVW*, 2017, pp. 1199–1209. DOI: 10.1109/ICCVW.2017.145.
9. Uzktent B., Ermon S. Learning when and where to zoom with deep reinforcement learning. *Proc. IEEE/CVF CVPR*, 2020, pp. 12342–12351. DOI: 10.1109/CVPR42600.2020.01236.
10. Mousavi H.K., Nazari M., Takáč M. et al. Multi-agent image classification via reinforcement learning. *Proc. IEEE/RSJ IROS*, 2019, pp. 5020–5027. DOI: 10.1109/IROS40897.2019.8968129.
11. Elizarov A.A., Razinkov E.V. Image classification using reinforcement learning. *RDLJ*, 2020, vol. 23, no. 6, pp. 1172–1191. DOI: 10.26907/1562-5419-2020-23-6-1172-1191 (in Russ.).
12. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proc. IEEE/CVF CVPR*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.
13. Girshick R. Fast R-CNN. *Proc. IEEE ICCV*, 2015, pp. 1440–1448. DOI: 10.1109/ICCV.2015.169.
14. *ImageNet Dataset*. Available at: <http://image-net.org> (accessed June 15, 2021).
15. *Papers with Code: Fine-Grained Image Classification*. Available at: <https://paperswithcode.com/task/fine-grained-image-classification> (accessed June 20, 2021).
16. Mingxing T., Quoc V.L. EfficientNet: Rethinking model scaling for convolutional neural networks. *ArXiv*, 2019, art. 1905.11946v5. Available at: <https://arxiv.org/pdf/1905.11946.pdf> (accessed June 15, 2021).
17. Abdolmaleki A., Springenberg J.T., Degraeve J. et al. Relative entropy regularized policy iteration. *ArXiv*, 2018, art. 1812.02256. Available at: <https://arxiv.org/abs/1812.02256> (accessed June 15, 2021).
18. Auer P., Cesa-Bianchi N., Fischer P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 2002, vol. 47, pp. 235–256. DOI: 10.1023/A:1013689704352.

Для цитирования

Елизаров А.А. Метод адаптивной классификации изображений с использованием обучения с подкреплением // Программные продукты и системы. 2022. Т. 35. № 1. С. 028–036. DOI: 10.15827/0236-235X.137.028-036.

For citation

Elizarov A.A. The adaptive image classification method using reinforcement learning. *Software & Systems*, 2022, vol. 35, no. 1, pp. 028–036 (in Russ.). DOI: 10.15827/0236-235X.137.028-036.