

ЧИСЛЕННЫЕ МЕТОДЫ И АНАЛИЗ ДАННЫХ

ТРИГОНОМЕТРИЧЕСКАЯ СИСТЕМА ФУНКЦИЙ В ПРОЕКЦИОННЫХ ОЦЕНКАХ ПЛОТНОСТИ ВЕРОЯТНОСТИ НЕЙРОСЕТЕВЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ

А.В. Савченко¹

¹Национальный исследовательский университет «Высшая школа экономики», Нижний Новгород, Россия

Аннотация

Исследована задача распознавания изображений, которые описываются векторами признаков высокой размерности, выделенными с помощью глубокой свёрточной нейронной сети и анализа главных компонент. Рассмотрена проблема высокой вычислительной сложности статистического подхода с непараметрическими оценками плотности вероятности векторов признаков, реализованного в вероятностной нейронной сети. Предложен новый метод статистической классификации на основе проекционных оценок плотности распределения с тригонометрической системой ортогональных функций. Показано, что такой подход позволяет преодолеть недостатки вероятностной нейронной сети, связанные с необходимостью обработки всех признаков всех эталонных изображений. В рамках экспериментального исследования для наборов изображений Caltech-101 и CASIA WebFaces показано, что предлагаемый подход позволяет на 1–5 % снизить вероятность ошибки распознавания и в 1,5–6 раз повысить вычислительную эффективность по сравнению с исходной вероятностной нейронной сетью для малых выборок эталонных изображений.

Ключевые слова: статистическое распознавание образов, обработка изображений, глубокие свёрточные нейронные сети, вероятностная нейронная сеть, проекционные оценки, распознавание лиц.

Цитирование: Савченко, А.В. Тригонометрическая система функций в проекционных оценках плотности вероятности нейросетевых признаков изображений / А.В. Савченко // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 149–158. – DOI: 10.18287/2412-6179-2018-42-1-149-158.

Введение

Большинство исследований в области распознавания изображений [1] сосредоточены на повышении точности, надёжности и вычислительной эффективности существующих решений за счёт применения новых архитектур глубоких свёрточных нейронных сетей (СНС) [2, 3, 4], новых классификаторов и их ансамблей [5, 6, 7], методов извлечения характерных признаков [8], алгоритмов приближённого поиска ближайшего соседа [9, 10] и пр. Несмотря на наличие большого числа хорошо зарекомендовавших себя подходов, основанных на технологиях глубокого обучения [1], интенсивность исследований в этом направлении не снижается. И связано это с тем, что современные методы распознавания характеризуются недостаточной точностью в вытекающих из потребностей прикладных исследований задачах, например, при наличии в базе данных малых выборок наблюдений (десятки эталонов для каждого класса) [11, 12], которые невозможно использовать для успешного обучения сложных нейросетевых структур [3, 4]. В таком случае наиболее часто используется перенос знаний или адаптация предметной области [7, 13], в котором глубокие СНС применяются не как метод классификации, а как способ извлечения признаков [8, 10]. Такая СНС предварительно обучается с помощью большого набора изображений, например, ImageNET [14].

В результате встаёт задача выбора наиболее подходящего метода классификации [15]. Достаточно перспективным здесь выглядит универсальный статистический подход [16, 17], в котором предполагается, что

векторы признаков изображений одного класса подчиняются одному закону распределения. Распространённой реализацией непараметрического подхода к оценке распределений является вероятностная нейронная сеть (ВНС) [18]. В ней для аппроксимации неизвестных плотностей вероятностей по обучающим выборкам используется конечная смесь гауссовских ядерных функций Розенблатта–Парзена. ВНС характеризуется чрезвычайно быстрой процедурой обучения и сходимостью к оптимальному байесовскому решению. Серьёзным недостатком аппарата ВНС [19, 20] является требование к хранению и обработке всех элементов всех обучающих выборок (*memory-based approach*), что делает её физически не реализуемой для сложных приложений, таких как распознавание изображений при наличии жёстких ограничений на время распознавания и/или на объём оперативной памяти, особенно для большого числа классов [9, 17].

Поэтому актуальной становится задача синтеза непараметрических статистических классификаторов, не требующих хранения и сопоставления всех признаков. Для её решения в настоящей работе показано, что если воспользоваться предположением о независимости главных компонент, выделенных из вектора признаков изображений [15], а для оценки плотности вероятности каждой компоненты вместо ядра Розенблатта–Парзена использовать проекционные оценки [21, 22, 23] на основе тригонометрических функций [24, 25], то возможно преодолеть отмеченные недостатки ВНС без потерь в скорости обучения. При этом получившаяся модификация ВНС будет также сходиться к

оптимальному байесовскому решению. В ходе экспериментального исследования в задаче классификации категорий изображений и распознавания лиц для популярных баз данных Caltech-101 [26] и CASIA Web-Faces [27] продемонстрировано повышение точности и вычислительной эффективности распознавания по сравнению с традиционными подходами в случае наличия ограниченного количества эталонов для каждого класса.

1. Задача распознавания изображений в условиях малого числа наблюдений на основе вероятностной нейронной сети

Задача распознавания состоит в том, чтобы поступающему на вход изображению одного объекта поставить в соответствие один из $C > 1$ классов [1, 12]. При этом для каждого c -го класса доступен набор из $R(c) \geq 1$ эталонных изображений. Рассмотрим далее случай *малых выборок* [6, 12, 21]: $R \approx 1$, который характерен для многих систем обработки изображений. Для каждого доступного изображения осуществляется извлечение характерных признаков. В наиболее часто используемых сейчас методах переноса знаний [2, 13] для настройки классификатора может использоваться не доступное обучающее множество, а внешняя база данных изображений, с помощью которой происходит обучение глубокой СНС [3, 4]. Далее для распознавания произвольных изображений они приводятся к одному размеру (высота U и ширина V) и подаются на вход СНС [8, 10]. Выходы из $D \gg 1$ значений одного предпоследнего слоя нейронной сети нормируются (в метрике L_2) и формируют вектор признаков \mathbf{x} этого изображения с размерностью D . Аналогичная процедура применяется для извлечения D -мерного вектора признаков $\mathbf{x}_r(c)$ из каждого r -го эталонного изображения c -го класса [12].

В итоге на этапе распознавания обучается классификатор выделенных признаков. В традиционном переносе знаний применяется логистическая регрессия – последний полносвязный слой исходной СНС заменяется на новый слой с C выходами (по одному на каждый класс исходной задачи), и происходит дообучение (*fine-tuning*) полученной нейросети для доступного обучающегося множества из R эталонов [2, 13].

В случаях малого числа наблюдений такая процедура оказывается недостаточно эффективной [11, 12], поэтому зачастую классификатор упрощается, например, с помощью снижения размерности векторов признаков [28] на основе анализа главных компонент [15], при котором векторы \mathbf{x} и $\mathbf{x}_r(c)$ линейно преобразовываются в векторы $M \ll D$ главных компонент $\mathbf{t} = [t_1, \dots, t_M]$ и $\mathbf{t}_r(c) = [t_{r,1}(c), \dots, t_{r,M}(c)]$ соответственно. Далее могут применяться известные методы построения небольшого числа опорных подпространств [29] или отбора наиболее информативных эталонов [19, 30].

Проблема состоит в том [1, 17], что каждому конкретному образу обычно присуща известная вариативность, т.е. изменчивость его признаков от одного образца наблюдения к другому, которая носит

случайный характер. Обычно преодоление данной проблемы связывают со статистическим подходом [15, 16], когда в роли образа выступает соответствующий закон распределения \mathbf{P}_c векторов признаков объектов одного класса. В таком случае задача сводится к проверке C гипотез W_c о законе распределения \mathbf{P} признаков входного изображения:

$$W_c : \mathbf{P} = \mathbf{P}_c, c = \overline{1, C}. \quad (1)$$

Её оптимальное решение дает байесовский критерий минимума среднего риска [16] – делается вывод о принадлежности входного объекта к классу с максимальной апостериорной вероятностью

$$c^* = \operatorname{argmax}_{c \in \{1, \dots, C\}} f(\mathbf{t} | W_c) \cdot P(W_c). \quad (2)$$

Здесь $P(W_c)$ – априорная вероятность появления c -го класса, $f(\mathbf{t} | W_c)$ – условная плотность вероятности (оценка распределения \mathbf{P}_c) главных компонент векторов признаков класса c . Для оценки априорной вероятности зачастую [18] используется соотношение числа эталонов в обучающих выборках $P(W_c) = R(c)/R$, где $R = \sum_{c=1}^C R(c)$ – общее число эталонов. Восстановление неизвестных законов распределения \mathbf{P}_c происходит в процессе предварительного обучения по выборкам $\{\mathbf{t}_r(c)\}$ конечных объёмов $R(c)$. В предположении о нормальном распределении \mathbf{P}_c могут применяться модификации линейного дискриминантного анализа [11]. С точки зрения современной прикладной статистики на практике обычно нет оснований полагать справедливость предположения о нормальном законе распределения для произвольных объектов [31, 32]. Поэтому большей популярностью в настоящее время пользуются *непараметрические* методы оценивания распределений [33], например [18]:

$$\hat{f}(\mathbf{t} | W_c) = \frac{1}{R(c)} \sum_{r=1}^{R(c)} K(\mathbf{t}, \mathbf{t}_r(c)) \quad (3)$$

на основе гауссовского ядра Розенблатта–Парзена

$$K(\mathbf{t}, \mathbf{t}_r(c)) = \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{m=1}^M (t_m - t_{r,m}(c))^2\right). \quad (4)$$

Здесь σ – фиксированный параметр сглаживания. С учётом оценки (3) итоговое решение (2) может быть записано в виде

$$c^* = \operatorname{argmax}_{c \in \{1, \dots, C\}} \frac{1}{R} \sum_{r=1}^{R(c)} K(\mathbf{t}, \mathbf{t}_r(c)). \quad (5)$$

Критерий (5) и представляет собой реализацию ВНС [18]. Здесь для каждого класса происходит сопоставление входного объекта со *всеми* признаками $\mathbf{t}_r(c)$ *каждого* эталона. В таком случае алгоритмическая сложность реализации критерия (5) оценивается как $O(RM)$. Сложность по затратам памяти $O(RM)$ оказывается высокой из-за необходимости хранения

всех эталонных объектов. В результате применение ВНС может быть неприемлемо для многих прикладных систем, функционирующих на малопроизводительном оборудовании [9].

Так как известного повышения вычислительной эффективности ВНС за счёт использования подходящих структур данных [34] оказывается обычно недостаточно, преодоление указанной проблемы связывают с выделением в обучающем множестве информативных эталонов [19, 30], которые и участвуют в дальнейшем распознавании. Наиболее популярный подход [19] сводится к предварительной кластеризации обучающего множества (чаще всего, на основе метода *k-means* [15]) и применении центроидов выделенных кластеров для обучения ВНС (3). К сожалению, такое решение приводит к потере основных достоинств ВНС. Действительно, процедура обучения становится достаточно сложной, а само решение перестаёт быть оптимальным в байесовском смысле, особенно для существенно различающихся по объёму кластеров [35].

2. Проекционные оценки плотности вероятности в вероятностной нейронной сети

В настоящем параграфе оценки плотности вероятности [36, 37] с помощью гауссовских ядерных функций Розенблатта–Парзена [18] заменяются на известные проекционные оценки [38, 39], в которых плотность вероятности записывается как сумма ортогональных разложений [40, 41]. Ситуация резко усложняется для многомерного случая, в котором образуется многомерная ортонормированная система из всевозможных произведений одномерных базисных функций. В результате с ростом размерности M вектора признаков экспоненциально возрастает сложность вычисления ортонормированных функций [36] и, как следствие, требование к минимальному числу эталонов каждого класса, достаточных для оценки многомерного распределения. Поэтому использование такого подхода для классификации векторов значений признаков большой размерности на практике в общем случае не представляется возможным. В то же время стоит отметить, что для рассматриваемого случая – признаков главных компонент, выделенных из выходов глубокой СНС, – можно использовать предположение о независимости отдельных компонент векторов \mathbf{t} и $\mathbf{t}_r(c)$ [35]. Тогда плотность вероятности каждого класса запишется как

$$\hat{f}(\mathbf{t}|W_c) = \prod_{m=1}^M f_c(t_m), \quad (6)$$

где одномерную плотность вероятности i -го признака $f_c(t_m)$ можно оценить аналогично (3). Если вместо ядра Розенблатта–Парзена (4) воспользоваться проекционными оценками на основе ортогональной системы тригонометрических функций [24, 42]

$$\begin{cases} \Psi_j^{(1)}(t) = \cos(\pi j t), \\ \Psi_j^{(2)}(t) = \sin(\pi j t) \end{cases} \quad (7)$$

и выполнить отсечение первых J членов тригонометрического ряда, то оценка плотности вероятности в (6) примет следующий вид [43]:

$$\hat{f}_c(t_m) = \frac{1}{R(c)} \sum_{r=1}^{R(c)} \sum_{j=0}^J \sum_{k=1}^2 \Psi_j^{(k)}(t_m) \Psi_j^{(k)}(t_{r,m}(c)). \quad (8)$$

К сожалению, реализация такого подхода оказывается в $2J$ раз медленнее по сравнению с традиционной ВНС (5). Однако если выполнить несложные тригонометрические преобразования выражения (8), то можно получить эквивалентную оценку плотности вероятности с помощью ядра Дирихле [22, 40]:

$$\hat{f}_c(t_m) = \frac{1}{2R(c)} \sum_{r=1}^{R(c)} \frac{\sin\left(\left(J + \frac{1}{2}\right)(t_m - t_{r,m}(c))\right)}{2 \sin(t_m - t_{r,m}(c))}. \quad (9)$$

В результате асимптотическая сложность классификации (2), (6), (9) будет совпадать со сложностью ВНС. Именно в таком виде (с учетом различий в применяемой ядерной функции) традиционно записывается ВНС на основе проекционных оценок [36, 37], которая, хоть и приводит в ряде случаев к повышению точности, очевидно, не влияет на вычислительную эффективность классификации.

3. Предложенный подход на основе проекционных оценок плотности

Стоит отметить, что для тригонометрической системы функций не гарантируется неотрицательное значение выражения (9) [22, 43]. В связи с этим на практике обычно используются другие ортогональные функции, такие как многочлены Лагерра или Лежандра [22]. В нашей работе реализован альтернативный подход, в котором для оценки плотности вероятности $f_c(t_m)$ вычисляется среднее арифметическое первых J частичных сумм тригонометрического ряда [43]. Эта оценка представима в виде (3), где вместо ядра Розенблатта–Парзена применяется ядро Фейера [40, 41]

$$\hat{f}_c(t_m) = \frac{1}{R(c)} \frac{1}{2(J+1)} \times \sum_{r=1}^{R(c)} \left(\frac{\sin\left(\frac{J+1}{2}(t_m - t_{r,m}(c))\right)}{\sin\left(\frac{1}{2}(t_m - t_{r,m}(c))\right)} \right)^2. \quad (10)$$

Такой подход, хоть и приводит к неотрицательным значениям оценки плотности вероятности (6), не позволяет повысить вычислительную эффективность распознавания [20, 43]. Поэтому в настоящей работе предлагается воспользоваться другим, эквивалентным (10) выражением [35, 43]:

$$\hat{f}_c(t_m) = 0,5 + \sum_{j=1}^J \sum_{k=1}^2 a_{m;j}^{(k)}(c) \cdot \Psi_j^{(k)}(t_m). \quad (11)$$

Здесь коэффициенты ряда предложено оценить по имеющейся обучающей выборке как среднее арифме-

тическое $(J+1)$ обычных коэффициентов тригонометрического ряда (7), где $m = \overline{1, M}, j = \overline{0, J}, k = \overline{1, 2}$:

$$a_{m;j}^{(k)}(c) = \frac{J-j+1}{R(c) \cdot (J+1)} \sum_{r=1}^{R(c)} \psi_j^{(k)}(t_{r,m}(c)). \quad (12)$$

Тогда итоговое выражение для распознавания изображений (2), (6) при переходе к более удобному логарифму правдоподобия запишется в виде

$$\max_{c \in \{1, \dots, C\}} \left(\log R(c) + \sum_{m=1}^M \log \left(0,5 + \sum_{j=1}^J \sum_{k=1}^2 a_{m;j}^{(k)}(c) \psi_j^{(k)}(t_m) \right) \right). \quad (13)$$

Таким образом, предлагаемый метод состоит в следующем. На этапе обучения для каждого класса вычисляются $M(2J+1)$ коэффициентов с помощью процедуры (12), имеющей линейную сложность. Далее в процессе распознавания изображение подаётся на вход СНС, выход предпоследнего слоя нормируется и преобразуется в последовательность из M главных компонент. После этого для каждой компоненты вычисляются J значений базисных функций (7). Для ускорения этой процедуры можно воспользоваться известными рекурсивными выражениями для тригонометрических функций суммы переменных:

$$\begin{aligned} \psi_j^{(1)}(t) &= \psi_{j-1}^{(1)}(t) \cdot \psi_1^{(1)}(t) - \psi_{j-1}^{(2)}(t) \cdot \psi_1^{(2)}(t), \\ \psi_j^{(2)}(t) &= \psi_{j-1}^{(1)}(t) \cdot \psi_1^{(2)}(t) + \psi_{j-1}^{(2)}(t) \cdot \psi_1^{(1)}(t) \end{aligned} \quad (14)$$

с инициализацией $\psi_1^{(1)}(t) = \cos(\pi t), \psi_1^{(2)}(t) = \sin(\pi t)$. В итоге количество сложных операций вычисления тригонометрических функций сокращается в J раз (по сравнению с (7)).

Далее для каждого класса и каждой компоненты вектора \mathbf{t} оценивается логарифм правдоподобия (13). Итоговое решение принимается в пользу максимально правдоподобного класса.

Остановимся подробнее на преимуществах предложенного подхода. Во-первых, сохраняются все основные преимущества традиционной ВНС: сходимость к байесовскому решению и чрезвычайно быстрая процедура обучения. Стоит отметить, что в предложенном подходе можно эффективно выполнить до обучение при появлении новых эталонных изображений. Действительно, при появлении нового эталона $\mathbf{t}_{R(c)+1}(c)$ коэффициенты (12) могут быть скорректированы за константное время следующим образом

$$\tilde{a}_{m;j}^{(k)}(c) = \frac{R(c) \cdot a_{m;j}^{(k)}(c) + \psi_j^{(k)}(t_{R(c)+1,m}(c))}{R(c)+1}. \quad (15)$$

Наконец, основным преимуществом предлагаемого подхода является низкая алгоритмическая сложность его реализации $O(CMJ)$. Как известно [22, 38, 44], сходимость ряда (11) обеспечивается в том случае, если параметр J определяется как $O(\sqrt[3]{R(c)})$. В результате вычислительная сложность предлагаемого алгоритма может быть оценена как $O\left(M \sum_{c=1}^C \sqrt[3]{R(c)}\right)$.

В худшем случае, когда для каждого класса доступен только $R(c) = 1$ эталон, асимптотическая сложность реализации разработанной модификации и исходной ВНС одинаковы: $O(MR)$. На практике в таком случае предложенный подход может оказаться в несколько раз медленнее за счёт суммирования в (13) $2J+1$ слагаемых. Наибольший выигрыш в скорости распознавания достигается для сбалансированных классов ($R(c) = R/C$). Тогда распознавание одного изображения в среднем оказывается приблизительно в $R(c) / (2\sqrt[3]{R(c)} + 1) \approx (R/C)^{2/3} / 2$ раз быстрее по сравнению с ВНС (3)–(5). Таким образом, предложенный алгоритм стоит применять вместо ВНС в случае $R(c) / (2\sqrt[3]{R(c)} + 1) \geq 1$, то есть при наличии в обучающем множестве в среднем не менее $R(c) = 5$ эталонных изображений. Кроме того, стоит отметить, что сложность по затратам памяти также снижается: $O(CD \cdot (R/C)^{1/3}) = O(D \cdot R^{1/3} \cdot C^{2/3})$ за счёт отказа от обработки всех элементов всех обучающих выборок (*memory-based approach*) и необходимости сохранения только коэффициентов (12).

Таким образом, разработанная модификация (12)–(15) позволяет преодолеть существующие недостатки (низкая вычислительная эффективность, отсутствие обобщения) оригинальной ВНС (4)–(5), сохранив при этом высокую скорость обучения и сходимость к оптимальному байесовскому решению. В следующем параграфе экспериментально продемонстрировано, что и для реальных задач распознавания изображений предложенный подход является не менее эффективным, чем традиционные классификаторы.

4. Результаты экспериментальных исследований

В настоящем параграфе описано сравнение разработанного метода (12)–(14) с оригинальной ВНС, методом k -ближайших соседей (k -БС) для метрики Евклида и реализациями классификаторов из библиотеки OpenCV: машина опорных векторов (SVM), random forest и многослойный персептрон (искусственная нейронная сеть прямого распространения). Кроме того, для классификаторов, основанных на полном переборе множества эталонов (k -БС и ВНС), использовался известный способ повышения вычислительной эффективности [19] за счёт предварительной кластеризации (с помощью алгоритма k -medoids) каждого эталонного класса и дальнейшего сравнения входного изображения только с выделенными центрами кластеров (медоидами) [15]. Предложенный подход и все остальные методы были реализованы на языке C++ в специальном приложении [45]. В наших экспериментах применялись следующие параметры классификаторов: кластеризация с $k = 10$ медоидами; one-versus-all SVM с коэффициентом регуляризации $C = 1$ и гауссовским ядром; random forest из 100 деревьев; многослойный персептрон с 256 нейронами на скрытом слое, функцией активации типа сигмоида, алгоритмом обучения g-pror, на вход которому поступает 256 главных компонент. Приведённые параметры обеспе-

чивают достаточно низкое смещение классификаторов, что позволило достичь наилучшей точности для всех анализируемых размеров выборок и наборов данных. Однако в методе k -ближайших соседей при наличии не более 10 эталонов каждого класса параметр k устанавливался равным 1, а для большего числа эталонов $k = 3$.

Эксперименты проводились на ноутбуке MacBook Pro 2015 (16 Гб ОЗУ, 4-ядерный процессор Intel Core i7 2,2 ГГц). В первом эксперименте рассмотрено применение предложенной модификации ВНС в задаче распознавания $C=101$ категорий из набора Caltech-101, содержащего 8677 изображений [26]. Для извлечения признаков использовалась библиотека Caffe [46] и две глубоких СНС – Inception v1 (GoogLeNet) [3] и VGGNet с 19 слоями [4]. Нейросетевые модели, предварительно обученные для распознавания 1000 классов изображений ImageNet, были загружены с официального репозитория Caffe Model Zoo. На вход СНС подавались матрицы цветных (RGB) изображений из набора Caltech-101, приведённых к одной размерности $U=V=224$ пикселя. Выходы слоев «pool5/7×7-s1» (Inception v1) и «fc6» (VGGNet-19) нормировались в метрике L_2 , после чего извлекались все главные компоненты для получения окончательных векторов признаков размерности $M=1024$ и $M=4096$ для Inception и VGGNet-19 соответственно.

В течение 10 раз повторялся следующий эксперимент. В обучающее множество из каждого класса наугад выбиралось фиксированное число изображений $R(c)$, а тестирование проводилось на всех остальных изображениях. Для проверки значимости различий в точности и времени распознавания применялся критерий Мак-Немара с уровнем значимости 0,05. Оценки точности классификации α (%) и среднего времени распознавания одного изображения t (мс) для 10 и 25 эталонных изображений в каждом классе $R(c)$ представлены в табл. 1 и табл. 2 для признаков, извлечённых с помощью Inception и VGGNet-19 соответственно.

Табл. 1. Результаты распознавания изображений для набора фотографий Caltech-101, Inception v1

| | $R(c) = 10$ | | $R(c) = 25$ | |
|--|--------------|----------------|--------------|----------------|
| | $\alpha, \%$ | $t, \text{мс}$ | $\alpha, \%$ | $t, \text{мс}$ |
| к-БС | 74,12 | 1,93 | 79,35 | 5,66 |
| к-БС, k-medoids | 73,78 | 1,08 | 76,90 | 1,10 |
| Random Forest | 51,98 | 0,37 | 67,12 | 0,41 |
| Многослойный перцептрон | 59,65 | 0,17 | 81,31 | 0,17 |
| SVM | 65,46 | 1,17 | 83,83 | 2,96 |
| ВНС | 74,61 | 1,89 | 80,27 | 5,51 |
| ВНС, k-medoids | 74,91 | 1,01 | 77,78 | 1,07 |
| Предложенный подход (12) – (15) | 78,06 | 1,14 | 82,30 | 1,49 |

Здесь извлечение признаков с помощью VGGNet-19 привело к меньшей точности распознавания по сравнению с GoogLeNet для малого числа эталонных

изображений каждого класса ($R(c)=10$), но при повышении числа эталонов для ряда методов классификации признаки VGGNet оказались более точными. Подтвердился известный факт о недостаточной эффективности сложных классификаторов (SVM, random forest, искусственная нейронная сеть) при небольших объемах обучающих выборок.

Табл. 2. Результаты распознавания изображений для набора фотографий Caltech-101, VGGNet-19

| | $R(c) = 10$ | | $R(c) = 25$ | |
|--|--------------|----------------|--------------|----------------|
| | $\alpha, \%$ | $t, \text{мс}$ | $\alpha, \%$ | $t, \text{мс}$ |
| к-БС | 70,99 | 7,40 | 80,46 | 21,56 |
| к-БС, k-medoids | 69,61 | 4,06 | 74,26 | 4,33 |
| Random Forest | 50,14 | 0,29 | 70,57 | 0,42 |
| Многослойный перцептрон | 71,69 | 0,79 | 84,36 | 0,85 |
| SVM | 64,96 | 4,22 | 83,62 | 11,03 |
| ВНС | 69,92 | 7,37 | 79,14 | 21,31 |
| ВНС, k-medoids | 69,68 | 3,89 | 74,38 | 4,14 |
| Предложенный подход (12) – (15) | 74,77 | 5,17 | 82,53 | 6,14 |

В то же время среднее время распознавания t для этих методов (за исключением SVM) оказалось наиболее малым среди всех рассматриваемых подходов. Время распознавания изображения для машины опорных векторов является достаточно высоким в связи с реализацией для каждого класса отдельного бинарного классификатора в режиме one-versus-all. В методах, основанных на полном переборе обучающего множества (к-БС и ВНС), время классификации линейно возрастает при увеличении числа эталонов R . А применение предварительной кластеризации (k-medoids), хоть и позволяет в 1,5–4,5 раза ускорить распознавание, приводит к снижению точности на 0,5–6%. Наконец, основной вывод по результатам эксперимента состоит в том, что точность предложенного подхода (12)–(15) оказалась на 2–5% выше точности оригинальной ВНС, а среднее время распознавания снизилось в 1,5–3,5 раза. Выигрыши в точности и вычислительной эффективности по сравнению с ВНС во всех случаях оказались статистически значимыми. При этом во всех наших экспериментах с этим набором данных для достаточно малого числа эталонов в каждом классе ($R(c) \leq 20$) предложенный подход оказался наиболее точным среди всех классификаторов, хотя при повышении размера обучающей выборки традиционные методы ожидаемо становятся более точными.

В следующем эксперименте рассмотрена задача идентификации лиц [1, 12]. Использовались 66000 фотографий первых $C=1000$ людей из базы данных фотографий лиц CASIA WebFaces [27]. Для извлечения признаков применялась наиболее точная из свободно доступных нейросетевых моделей – Light СНС (версия С) [47], обученная её авторами с помощью сверхбольшого набора фотографий лиц MS-Celeb-1M. Эта СНС извлекает $D=256$ вещественных признаков из полутонного изображения лица с высотой

$U=128$ и шириной $V=128$ пикселей. Результаты эксперимента приведены в табл. 3.

Табл. 3. Результаты распознавания лиц для набора фотографий CASIA WebFaces, Light SNC

| | $R(c) = 15$ | | $R(c) = 30$ | |
|--------------------------------------|--------------|----------------|--------------|----------------|
| | $\alpha, \%$ | $t, \text{мс}$ | $\alpha, \%$ | $t, \text{мс}$ |
| к-БС | 73,36 | 9,71 | 78,31 | 27,48 |
| к-БС, k-medoids | 68,92 | 4,84 | 75,12 | 4,95 |
| Random Forest | 67,47 | 1,34 | 76,23 | 1,43 |
| Многослойный перцептрон | 72,38 | 1,95 | 80,47 | 1,96 |
| SVM | 67,27 | 3,68 | 77,46 | 4,95 |
| ВНС | 75,05 | 8,42 | 77,62 | 22,73 |
| ВНС, k-medoids | 70,83 | 2,93 | 74,85 | 3,22 |
| Предложенный подход (12)–(15) | 77,07 | 3,14 | 81,94 | 3,88 |

Здесь подтвердились основные выводы предыдущего эксперимента: высокая вычислительная сложность ВНС и к-БС препятствует их практической реализации, а предварительный отбор эталонов с помощью k-medoids приводит к значимому повышению вероятности ошибочной классификации на 3–5%.

В то же время предложенная модификация (12)–(15) во всех случаях является более предпочтительной как по вычислительной эффективности (в 2,7–6 раз), так и по точности (на 2–4%), чем оригинальная ВНС (4)–(5). При этом с точки зрения критерия Мак-Немара повышение эффективности по сравнению с ВНС оказывается статистически значимым. Стоит отметить, что в связи с большим числом классов разработанный подход даже для $R(c)=30$ эталонов оказывается более точным по сравнению с остальными методами классификации.

Заключение

Таким образом, в настоящей работе предложена модификация ВНС (12)–(15), основанная на проекционных оценках плотности вероятности, которые используют систему тригонометрических функций и предположение о независимости признаков классифицируемого объекта. Показано, что разработанная модификация сохраняет основные преимущества оригинальной ВНС: сходимость к байесовскому решению, линейное время обучения и константная сложность дообучения. При этом предложенный подход является намного более эффективным с точки зрения вычислительной сложности и затрат памяти (выигрыш до $(R/C)^{2/3}/2$ раз) за счёт отказа от хранения признаков всех эталонов и обучения модели с помощью вычисления коэффициентов (12) усреднённого тригонометрического ряда (11). Результаты экспериментального исследования для признаков, извлечённых с помощью современных СНС Inception, VGGNet и Light SNC, показали, что разработанная модификация не только во всех случаях оказывается предпочтительнее исходной ВНС, но и является наиболее точным классификатором для малого числа эталонных изображений.

Основным недостатком предложенного подхода по сравнению с классической ВНС является отсутствие в нём аналога выражения (4), в котором вычисляется расстояние между векторами признаков входного и эталонного изображения. В результате для нашей модификации оказывается недоступным повышение точности ВНС за счёт применения в (4) мер близости более сложных, чем метрика Евклида [5]. Кроме того, стоит отметить, что реализация предложенного подхода оказалась медленнее оригинальной ВНС при наличии очень малого (менее 5) числа эталонов для каждого класса в связи с оценкой отдельной плотности вероятности для каждого признака при вычислении выражения (6).

В то же время следует отметить необходимость проведения ряда дополнительных исследований. Прежде всего, следует оценить точность распознавания изображений для оценки плотности вероятности с помощью других известных ортогональных систем (например, многочленов Лагерра или Лежандра) [22, 23]. Также стоит проанализировать возможность адаптации разработанного метода к обработке больших данных, например, на основе последовательного анализа [48, 49] иерархического представления изображений [2] с применением предложенного подхода для предварительной отбраковки большинства классов и выбора окончательного решения с помощью одного из традиционных классификаторов. Наконец, представляет интерес исследование применимости предложенного подхода для признаков, выделенных СНС с небольшим числом параметров, таких как MobileNet, SqueezeNet, в расчёте на их реализацию на мобильных устройствах.

Благодарности

Исследование выполнено при поддержке гранта президента РФ для молодых ученых – докторов наук № МД-306.2017 и Лаборатории алгоритмов и технологий анализа сетевых структур (ЛАТАС) Национального исследовательского университета Высшая школа экономики. Работа параграфов 3 и 4 выполнена за счёт гранта Российского научного фонда (проект № 14-41-00039).

Литература

1. **Prince, S.J.D.** Computer vision: Models, learning, and inference / S.J.D. Prince. – Cambridge: Cambridge University Press, 2012. – 598 p. – ISBN: 978-1-107-01179-1.
2. **Goodfellow, I.** Deep learning / I. Goodfellow, Y. Bengio, A. Courville. – Cambridge, London: The MIT Press, 2016. – 800 p. – ISBN: 9780262035613.
3. **Szegedy, C.** Going deeper with convolutions / C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich // Proceedings of the 2015 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 1-9. – DOI: 10.1109/CVPR.2015.7298594.
4. **Simonyan, K.** Very deep convolutional networks for large-scale image recognition / K. Simonyan, A. Zisserman // arXiv preprint arXiv:1409.1556, 2014.

5. **Savchenko, A.V.** Probabilistic neural network with homogeneity testing in recognition of discrete patterns set / A.V. Savchenko // *Neural Networks*. – 2013. – Vol. 46. – P. 227-241. – DOI: 10.1016/j.neunet.2013.06.003.
6. **Krizhevsky, A.** ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS'12)*. – 2012. – Vol. 1. – P. 1097-1105.
7. **Rassadin, A.G.** Group-level emotion recognition using transfer learning from face identification / A.G. Rassadin, A.S. Gruzdev, A.V. Savchenko // *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI)*. – 2017. – P. 544-548. – DOI: 10.1145/3136755.3143007.
8. **Sharif Razavian, A.** CNN features off-the-shelf: an astounding baseline for recognition / A. Sharif Razavian, H. Azizpour, J. Sullivan, S. Carlsson // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '14)*. – 2014. – P. 806-813. – DOI: 10.1109/CVPRW.2014.131.
9. **Savchenko, A.V.** Maximum-likelihood approximate nearest neighbor method in real-time image recognition / A.V. Savchenko // *Pattern Recognition*. – 2017. – Vol. 61. – P. 459-469. – DOI: 10.1016/j.patcog.2016.08.015.
10. **Savchenko, A.V.** Deep neural networks and maximum likelihood search for approximate nearest neighbor in video-based image recognition / A.V. Savchenko // *Optical Memory and Neural Networks (Information Optics)*. – 2017. – Vol. 26, Issue 2. – P. 129-136. – DOI: 10.3103/S1060992X17020102.
11. **Raudys, S.J.** Small sample size effects in statistical pattern recognition: Recommendations for practitioners / S.J. Raudys, A.K. Jain // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1991. – Vol. 13, Issue 3. – P. 252-264. – DOI: 10.1109/34.75512.
12. **Савченко, А.В.** Метод максимально правдоподобных рассогласований в задаче распознавания изображений на основе глубоких нейронных сетей / А.В. Савченко // *Компьютерная оптика*. – 2017. – Т. 41, № 3. – С. 422-430. – DOI: 10.18287/2412-6179-2017-41-3-422-430.
13. **Pan, S.J.** A survey on transfer learning / S.J. Pan, Q. Yang // *IEEE Transactions on Knowledge and Data Engineering*. – 2010. – Vol. 22, Issue 10. – P. 1345-1359. – DOI: 10.1109/TKDE.2009.191.
14. **Russakovsky, O.** ImageNet large scale visual recognition challenge / O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, F.-F. Li // *International Journal of Computer Vision*. – 2015. – Vol. 115, Issue 3. – P. 211-252. – DOI: 10.1007/s11263-015-0816-y.
15. **Theodoridis, S.** *Pattern recognition* / S. Theodoridis, C. Koutroubas. – 4th ed. – Burlington, San Diego, London: Elsevier Inc., 2009. – 840 p. – ISBN: 978-1-59749-272-0.
16. **Webb, A.R.** *Statistical pattern recognition* / A.R. Webb. – 2nd ed. – Chichester, England: John Wiley & Sons, Ltd., 2002. – ISBN: 978-0-470-84513-4.
17. **Савченко, А.В.** Распознавание изображений на основе вероятностной нейронной сети с проверкой однородности / А.В. Савченко // *Компьютерная оптика*. – 2013. – Т. 37, № 2. – С. 254-262. – ISSN 0134-2452.
18. **Specht, D.F.** Probabilistic neural networks / D.F. Specht // *Neural Networks*. – 1990. – Vol. 3, Issue 1. – P. 109-118. – DOI: 10.1016/0893-6080(90)90049-Q.
19. **Kusy, M.** Probabilistic neural network structure reduction for medical data classification / M. Kusy, J. Kluska // *Proceedings of International Conference on Artificial Intelligence and Soft Computing (ICAISC)*. – 2013. – P. 118-129. – DOI: 10.1007/978-3-642-38658-9_11.
20. **Savchenko, A.V.** Pattern classification with the probabilistic neural networks based on orthogonal series kernel / A.V. Savchenko // *Proceedings of International Symposium on Neural Networks (ISNN 2016)*. – 2016. – P. 505-512. – DOI: 10.1007/978-3-319-40663-3_58.
21. **Čencov, N.N.** *Statistical decision rules and optimal inference* / N.N. Čencov. – Providence, RI: American Mathematical Society, 2000. – ISBN: 978-0-8218-1347-8.
22. **Деврой, Л.** Непараметрическое оценивание плотности. L1-подход / Л. Деврой, Л. Дьёрфи. – М.: Мир, 1988. – 408 с. – ISBN: 5-03-000475-0.
23. **Efromovich, S.** *Nonparametric curve estimation: Methods, theory, and applications* / S. Efromovich. – New York: Springer, 1999. – ISBN: 978-0-387-98740-8.
24. **Greblicki, W.** Classification using the Fourier series estimate of multivariate density functions / W. Greblicki, M. Pawlak // *IEEE Transactions on Systems, Man, and Cybernetics*. – 1981. – Vol. 11, Issue 10. – P. 726-730. – DOI: 10.1109/TSMC.1981.4308594.
25. **Rutkowski, L.** Sequential pattern recognition procedures derived from multiple Fourier series / L. Rutkowski // *Pattern Recognition Letters*. – 1988. – Vol. 8, Issue 4. – P. 213-216. – DOI: 10.1016/0167-8655(88)90027-X.
26. **Fei-Fei, L.** One-shot learning of object categories / L. Fei-Fei, R. Fergus, P. Perona // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2006. – Vol. 28, Issue 4. – P. 594-611. – DOI: 10.1109/TPAMI.2006.79.
27. **Yi, D.** Learning face representation from scratch / D. Yi, Z. Lei, S. Liao, S.Z. Li // *arXiv preprint arXiv:1411.7923*. – 2014.
28. **Wasikowski, M.** Combating the small sample class imbalance problem using feature selection / M. Wasikowski, X. Chen // *IEEE Transactions on Knowledge and Data Engineering*. – 2010. – Vol. 22, Issue 10. – P. 1388-1400. – DOI: 10.1109/TKDE.2009.187.
29. **Жердев, Д.А.** Распознавание объектов по диаграммам рассеяния электромагнитного излучения на основе метода опорных подпространств / Д.А. Жердев, Н.Л. Казанский, В.А. Фурсов // *Компьютерная оптика*. – 2014. – Т. 38, № 3. – С. 503-510.
30. **Савченко, В.В.** Принцип минимума информационного рассогласования в задаче спектрального анализа случайных временных рядов в условиях малых выборок наблюдений / В.В. Савченко // *Известия высших учебных заведений. Радиофизика*. – 2015. – Т. 58, № 5. – С. 415-422.
31. **Орлов, А.И.** Развитие математических методов исследования (2006-2015 гг.) / А.И. Орлов // *Заводская лаборатория. Диагностика материалов*. – 2017. – Т. 83, № 1-1. – С. 78-86.
32. **Shatskikh, S.Ya.** Normality assumption in statistical data analysis / S.Ya. Shatskikh, L.E. Melkumova // *CEUR Workshop Proceedings*. – 2016. – Vol. 1638. – P. 763-768. – DOI: 10.18287/1613-0073-2016-1638-763-768.
33. **Лапко, А.В.** Непараметрические модели распознавания образов в условиях малых выборок / А.В. Лапко, С.В. Ченцов, В.А. Лапко // *Автометрия*. – 1999. – № 6. – С. 105-113.
34. **Franti, P.** Fast and memory efficient implementation of the exact PNN / P. Franti, T. Kaukoranta, D.-F. Shen, K.-S. Chang // *IEEE Transactions on Image Processing*. – 2000. – Vol. 9, Issue 5. – P. 773-777. – DOI: 10.1109/83.841516.
35. **Савченко, А.В.** Об одном способе повышения вычислительной эффективности вероятностной нейронной сети в задаче распознавания образов на основе проекци-

- онных оценок / А.В. Савченко // Информационные системы и технологии. – 2015. – № 4(90). – С. 28-38.
36. **Rutkowski, L.** Adaptive probabilistic neural networks for pattern classification in time-varying environment / L. Rutkowski // IEEE Transactions on Neural Networks. – 2004. – Vol. 15, Issue 4. – P. 811-827. – DOI: 10.1109/TNN.2004.828757.
 37. **Duda, P.** On the Cesaro orthogonal series-type kernel probabilistic neural networks handling non-stationary noise / P. Duda, J.M. Zurada // Proceedings of the 9th International Conference on Parallel Processing and Applied Mathematics (LNCS). – 2012. – Vol. 7203, Pt. I. – P. 435-442. – DOI: 10.1007/978-3-642-31464-3_44.
 38. **Schwartz, S.C.** Estimation of probability density by an orthogonal series / S.C. Schwartz // The Annals of Mathematical Statistics. – 1967. – Vol. 38, Issue 4. – P. 1261-1265.
 39. **Efromovich, S.** Orthogonal series density estimation / S. Efromovich // Wiley Interdisciplinary Reviews: Computational Statistics. – 2010. – Vol. 2, Issue 4. – P. 467-476. – DOI: 10.1002/wics.97.
 40. **Фихтенгольц, Г.М.** Курс дифференциального и интегрального исчисления / Г.М. Фихтенгольц. – Т. 3. – М.: Физматлит, 2001. – 662 с.
 41. **Зорич, В.А.** Математический анализ / В.А. Зорич. – Ч. 2. – М.: Наука, 1984. – 640 с.
 42. **Hall, P.** On trigonometric series estimates of densities / P. Hall // Annals of Statistics. – 1981. – Vol. 9, Issue 3. – P. 683-685.
 43. **Новосёлов, А.А.** Параметризация моделей управляемых систем / А.А. Новосёлов // Вестник Сибирского государственного аэрокосмического университета. – 2010. – № 5. – С. 52-56.
 44. **Hart, J.D.** On the choice of a truncation point in Fourier series density estimation / J.D. Hart // Journal of Statistical Computation and Simulation. – 1985. – Vol. 21, Issue 2. – P. 95-116. – DOI: 10.1080/00949658508810808.
 45. Система распознавания изображений [Электронный ресурс]. – URL: https://github.com/HSE-asavchenko/HSE_FaceRec/tree/master/src/recognition_testing (дата обращения 01.12.2017).
 46. **Jia, Y.** Caffe: Convolutional architecture for fast feature embedding / Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, T. Darrell // Proceedings of the 22nd ACM International Conference on Multimedia. – 2014. – P. 675-678. – DOI: 10.1145/2647868.2654889.
 47. **Wu, X.** A light CNN for deep face representation with noisy labels / X. Wu, R. He, Z. Sun, T. Tan // arXiv preprint arXiv:1511.02683. – 2017.
 48. **Savchenko, A.V.** Fast multi-class recognition of piecewise regular objects based on sequential three-way decisions and granular computing / A.V. Savchenko // Knowledge-Based Systems. – 2016. – Vol. 91. – P. 250-260. – DOI: 10.1016/j.knsys.2015.09.021.
 49. **Savchenko, A.V.** Sequential three-way decisions in efficient classification of piecewise stationary speech signals / A.V. Savchenko // Proceedings of International Joint Conference on Rough Sets (IJCRS 2017). – 2017. – Part II. – P. 264-277. – DOI: 10.1007/978-3-319-60840-2_19.

Сведения об авторе

Савченко Андрей Владимирович, 1985 года рождения, в 2008 году окончил Нижегородский государственный технический университет им. Р.Е. Алексеева по специальности «Прикладная математика и информатика». В 2010 году защитил диссертацию на соискание учёной степени кандидата технических наук по специальности 05.13.18 «Математическое моделирование, численные методы и комплексы программ». В 2015 г. присвоено учёное звание доцента по специальности 05.13.18. В 2016 году присуждена учёная степень доктора технических наук по специальности 05.13.01 «Системный анализ, управление и обработка информации». В настоящее время работает профессором кафедры информационных систем и технологий и старшим научным сотрудником лаборатории алгоритмов и технологий анализа сетевых структур в Национальном исследовательском университете Высшая школа экономики – Нижний Новгород. Автор более 100 научных работ. Область научных интересов: обработка мультимедийной информации, распознавание образов. E-mail: avsavchenko@hse.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 1 декабря 2017 г. Окончательный вариант – 19 января 2018 г.

TRIGONOMETRIC SERIES IN ORTHOGONAL EXPANSIONS FOR DENSITY ESTIMATES OF DEEP IMAGE FEATURES

A.V. Savchenko¹

¹National Research University Higher School of Economics, Nizhny Novgorod, Russia

Abstract

In this paper we study image recognition tasks in which the images are described by high dimensional feature vectors extracted with deep convolutional neural networks and principal component analysis. In particular, we focus on the problem of high computational complexity of a statistical approach with non-parametric estimates of probability density implemented by the probabilistic neural network. We propose a novel statistical classification method based on the density estimators with orthogonal expansions using trigonometric series. It is shown that this approach makes it possible to overcome the drawbacks of the probabilistic neural network caused by the memory-based approach of instance-based learning. Our experimental study with Caltech-101 and CASIA WebFace datasets demonstrates that the proposed approach reduces the error rate by 1–5% and increases the computational speed by 1.5–6 times when compared to the original probabilistic neural network for small samples of reference images.

Keywords: statistical pattern recognition, image processing, deep convolutional neural networks, probabilistic neural network, orthogonal series estimates, unconstrained face identification.

Citation: Savchenko AV. Trigonometric series in orthogonal expansions for density estimates of deep image features. *Computer Optics* 2018; 42(1): 149-158. DOI: 10.18287/2412-6179-2018-42-1-149-158.

Acknowledgements: The work is supported by Russian Federation President grant no. МД-306.2017.9 and Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics. The research in Sections 3 and 4 was supported by RSF (Russian Science Foundation) project No. 14-41-00039.

References

- [1] Prince SJD. *Computer vision: Models, learning, and inference*. Cambridge: Cambridge University Press; 2012. ISBN: 978-1-107-01179-1.
- [2] Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge, London: The MIT Press; 2016. ISBN: 9780262035613.
- [3] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. *Proc CVPR 2015*: 1-9. DOI: 10.1109/CVPR.2015.7298594.
- [4] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* 2014.
- [5] Savchenko AV. Probabilistic neural network with homogeneity testing in recognition of discrete patterns set. *Neural Networks* 2013; 46: 227-241. DOI: 10.1016/j.neunet.2013.06.003.
- [6] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proc NIPS'12 2012*; 1: 1097-1105.
- [7] Rassadin A, Gruzdev A, Savchenko A. Group-level emotion recognition using transfer learning from face identification. *Proc ICMI 2017*: 544-548. DOI: 10.1145/3136755.3143007.
- [8] Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: an astounding baseline for recognition. *Proc CVPRW'14 2014*: 806-813. DOI: 10.1109/CVPRW.2014.131.
- [9] Savchenko AV. Maximum-likelihood approximate nearest neighbor method in real-time image recognition. *Pattern Recognition*, 2017; 61: 459-469. DOI: 10.1016/j.patcog.2016.08.015.
- [10] Savchenko AV. Deep neural networks and maximum likelihood search for approximate nearest neighbor in video-based image recognition. *Opt Mem Neural Networks* 2017; 26(2): 129-136. DOI: 10.3103/S1060992X17020102.
- [11] Raudys SJ, Jain AK. Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1991, 13(3): 252-264. DOI: 10.1109/34.75512.
- [12] Savchenko AV. Maximum-likelihood dissimilarities in image recognition with deep neural networks [In Russian]. *Computer Optics* 2017, 41(3): 422-430. DOI: 10.18287/2412-6179-2017-41-3-422-430.
- [13] Pan SJ, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* 2010, 22(10): 1345-1359. DOI: 10.1109/TKDE.2009.191.
- [14] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Li FF. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision* 2015; 115(3): 211-252. DOI: 10.1007/s11263-015-0816-y.
- [15] Theodoridis S, Koutroumbas C, (eds). *Pattern recognition*. 4th ed. Burlington, San Diego, London: Elsevier Inc; 2009. ISBN: 978-1-59749-272-0.
- [16] Webb AR. *Statistical pattern recognition*. 2nd ed. Chichester, England: John Wiley & Sons, Ltd; 2002. ISBN: 978-0-470-84513-4.
- [17] Savchenko AV. Image recognition on the basis of probabilistic neural network with homogeneity testing [in Russian]. *Computer optics* 2013; 37(2): 254-262.
- [18] Specht DF. Probabilistic neural networks. *Neural Networks* 1990; 3(1): 109-118. DOI: 10.1016/0893-6080(90)90049-Q.
- [19] Kusy M, Kluska J. Probabilistic neural network structure reduction for medical data classification. *Proc ICAISC 2013*: 118-129. DOI: 10.1007/978-3-642-38658-9_11.
- [20] Savchenko AV. Pattern classification with the probabilistic neural networks based on orthogonal series kernel. *Proc ISSN 2016*: 505-512. DOI: 10.1007/978-3-319-40663-3_58.
- [21] Čencov NN. *Statistical decision rules and optimal inference*. Providence, RI: American Mathematical Society; 2000. ISBN: 978-0-8218-1347-8.
- [22] Devroye L, Györfi L. *Nonparametric density estimation: The L1 view*. Wiley; 1985. ISBN: 978-0-4718-1646-1.
- [23] Efromovich S. *Nonparametric curve estimation: Methods, theory, and applications*. New York: Springer; 1999. ISBN: 978-0-387-98740-8.
- [24] Greblicki W, Pawlak M. Classification using the Fourier series estimate of multivariate density functions. *IEEE Transactions on Systems, Man, and Cybernetics* 1981; 11(10): 726-730. DOI: 10.1109/TSMC.1981.4308594.
- [25] Rutkowski L. Sequential pattern recognition procedures derived from multiple Fourier series. *Pattern Recognition Letters* 1988; 8(4): 213-216. DOI: 10.1016/0167-8655(88)90027-X.
- [26] Fei-Fei L, Fergus R, Perona P. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2006; 28(4): 594-611. DOI: 10.1109/TPAMI.2006.79.
- [27] Yi D, Lei Z, Liao S, Li SZ. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*, 2014.
- [28] Wasikowski M, Chen X. Combating the small sample class imbalance problem using feature selection. *IEEE Transactions on Knowledge and Data Engineering* 2010; 22(10): 1388-1400. DOI: 10.1109/TKDE.2009.187.
- [29] Zherdev DA, Kazanskiy NL, Fursov VA. Object recognition by the radar signatures of electromagnetic field scattering on base of support subspaces method [In Russian]. *Computer Optics* 2014; 38(3): 503-510.
- [30] Savchenko VV. The principle of the information-divergence minimum in the problem of spectral analysis of the random time series under the condition of small observation samples. *Radiophysics and Quantum Electronics* 2015; 58 (5): 373-379. DOI: 10.1007/s11141-015-9611-4.
- [31] Orlov AI. Development of the methods of mathematical research (2006-2015) [In Russian]. *Industrial Laboratory. Diagnostics of Materials* 2017; 83(1-1): 78-86.
- [32] Shatskikh SYa, Melkumova LE. Normality assumption in statistical data analysis. *CEUR Workshop Proceedings*

- 2016; 1638: 763-768. DOI: 10.18287/1613-0073-2016-1638-763-768.
- [33] Lapko AV, Lapko VA, Chentsov SV. Nonparametric models of pattern recognition under conditions of small samples. *Optoelectronics, Instrumentation and Data Processing* 1999; 6: 83-90.
- [34] Franti P, Kaukoranta T, Shen DF, Chang KS. Fast and memory efficient implementation of the exact PNN. *IEEE Transactions on Image Processing* 2000; 9(5): 773-777. DOI: 10.1109/83.841516
- [35] Savchenko AV. On improvement of computing efficiency of the probabilistic neural network in pattern recognition problem based on the orthogonal series estimates of class densities [In Russian]. *Information Systems and Technologies* 2015; 4: 28-38.
- [36] Rutkowski L. Adaptive probabilistic neural networks for pattern classification in time-varying environment. *IEEE Transactions on Neural Networks* 2004; 15(4): 811-827. DOI: 10.1109/TNN.2004.828757.
- [37] Duda P, Zurada JM. On the Cesaro orthogonal series-type kernel probabilistic neural networks handling non-stationary noise. *Proceedings of International Conference on Parallel Processing and Applied Mathematics (LNCS)* 2012; 7203(I): 435-442. DOI: 10.1007/978-3-642-31464-3_44.
- [38] Schwartz SC. Estimation of probability density by an orthogonal series. *The Annals of Mathematical Statistics* 1967: 1261-1265.
- [39] Efromovich S. Orthogonal series density estimation. *Wiley Interdisciplinary Reviews: Computational Statistics* 2010, 2(4): 467-476. DOI: 10.1002/wics.97.
- [40] Fichtenholz GM. Course of differential and integral calculus: Volume 3 [In Russian]. Moscow, "Fizmatlit" Publisher; 2001.
- [41] Zorich VA. *Mathematical analysis: Part 2* [In Russian]. Moscow, "Nauka" Publisher; 1984.
- [42] Hall P. On trigonometric series estimates of densities. *Annals of Statistics* 1981; 9(3): 683-685.
- [43] Novosyolov AA. Parametrization of models of control systems [In Russian]. *Herald of Siberian State Aerospace University* 2010; 5: 52-56.
- [44] Hart JD. On the choice of a truncation point in Fourier series density estimation. *Journal of Statistical Computation and Simulation* 1985; 21(2): 95-116. DOI: 10.1080/00949658508810808.
- [45] Video-based face recognition software. Source: https://github.com/HSE-asavchenko/HSE_FaceRec/tree/master/src/re-cognition_testing.
- [46] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Darrell T. Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM International Conference on Multimedia* 2014: 675-678. DOI: 10.1145/2647868.2654889.
- [47] Wu X, He R, Sun Z., Tan T. A light CNN for deep face representation with noisy labels. *arXiv preprint arXiv:1511.02683* 2017.
- [48] Savchenko AV. Fast multi-class recognition of piecewise regular objects based on sequential three-way decisions and granular computing. *Knowledge-Based Systems* 2016; 91: 250-260. DOI: 10.1016/j.knsys.2015.09.021.
- [49] Savchenko AV. Sequential three-way decisions in efficient classification of piecewise stationary speech signals. *International Joint Conference on Rough Sets (IJCRS)* 2017; II: 264-277. DOI: 10.1007/978-3-319-60840-2_19.

Author's information

Andrey Vladimirovich Savchenko (b. 1985) graduated from N. Novgorod State Technical University in 2002, majoring in Applied Mathematics and Informatics. He defended his PhD in Mathematical Modeling, Numeric Methods and Software Complexes in 2010. He received the Doctor of Science degree in System Analysis, Control and Information Processing in 2016. Currently he works as the professor of Information Systems and Technologies department and senior researcher of the laboratory of Algorithms and Technologies in Network Analysis in National Research University Higher School of Economics, Nizhny Novgorod. He is the co-author of more than 100 scientific papers. Research interests include multimedia processing and pattern recognition. E-mail: avsavchenko@hse.ru.

Received December 1, 2017. The final version – January 19, 2018.