

This paper describes the intelligent system for solving *Project Scheduling Problem*: the task is creating schedule of agricultural equipment for doing farm works. We report a novel method for searching optimal scheduling of agricultural operations. The proposed approach is based on genetic algorithm and multi-agent systems and permits to find appropriate solution, which minimizes operation costs.

УДК 004.932.75'1

Кузьмицкий Н.Н.

## АКТУАЛЬНЫЕ ВОПРОСЫ ИСПОЛЬЗОВАНИЯ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ И ИХ КОМИТЕТОВ В РАСПОЗНАВАНИИ ОБРАЗОВ ЦИФР

**Введение.** Основной целью распознавания образов цифр является создание машин, обладающих технологиями чтения визуальной числовой информации со скоростью и точностью не ниже человеческой. Насколько близки научные результаты к ее достижению?

С первых лет исследований в области OCR технологиям анализа изображений изолированных цифр уделяется повышенное внимание, ввиду многочисленности сфер их потенциального применения. С тех пор научным сообществом накоплен большой объем информации в форме моделей классификаторов, методов формирования признаков, баз маркированных данных и др., описанных в многочисленной литературе. Особенно активно в рамках данной проблематики развивались нейросетевые технологии, преимуществами которых являются: автоматическое извлечение признаков, устойчивость к зашумленным данным, возможность параллельной и эффективной аппаратной реализации и др. [1].

Так, в конце 1990-х на базе многослойного перцептрона (MLP) и алгоритма обратного распространения Яном Лекуном была разработана сверточная нейросетевая модель (СНС, CNN) [2], которая, по мнению ряда авторов, наилучшим образом подходит для решения визуальных задач анализа документов [3]. Им же была создана знаменитая база рукописных образов MNIST, являющаяся важной точкой отсчета при сравнении моделей классификаторов. Последние результаты в распознавании тестового MNIST (23 ошибки на 10000 примерах [4]) продемонстрировали перспективность применения CNN в достижении указанной выше главной цели. Однако открытым остается вопрос: какова эффективность CNN классификатора, обученного на одной базе, в распознавании образов других баз?

Анализ литературы выявил существенный недостаток в описании исследований, направленных на решение данного вопроса. Среди них можно выделить лишь статью Сивалда [5], в которой на примере трех баз рукописных образов (MNIST, USPS, DIGITS) и различных методик выделения признаков показана уязвимость классификаторов, основанных на моделях k-NN, SVM и частично CNN, в их способности переносить знания с одного контрольного множества на другое. Он назвал данную проблему "хрупкостью" (слабостью в ИИ терминологии) и она в отношении к CNN является главным объектом интереса описываемого исследования. При этом большое внимание в статье уделяется способности CNN распознавать образы цифр различных типов и увеличению эффективности переноса знаний CNN между различными базами данных.

**Регуляризация изображений изолированных цифр.** Активное развитие OCR технологий за последние 20 лет привело к появлению интеллектуальных продуктов, удовлетворяющих практическим потребностям: систем распознавания текста, обработки почтовых, банковских документов и др. [6]. Однако зачастую залогом эффективности их внедрений являлось сужение допустимой вариативности данных.

*Типы изображений изолированных цифр.* В зависимости от способа происхождения, можно выделить три наиболее значимые группы изображений изолированных цифр: машинописные, рукописные и синтезированные, примеры которых приведены на рисунке 1.

Источником изображений первой группы могут являться сканированные документы, журналы и др., содержащие текстовую информацию шрифтовой формы. Ее особенностью является соблюдение правил написания отдельных образов и, как следствие, их низкая вариативность, что позволило создать весьма эффективные методики распознавания. Изображения второй группы, ввиду разнообразия стилей написания (наклона, ширины пера и др.), создают серьезные проблемы в решении классификационных задач (выборе признаков, параметров методов обучения и др.), которые не позволили до настоящего момента создать универсальную технологию распознавания, при этом имеются методы, демонстрирующие точность, сравнимую с человеческой на отдельных тестовых множествах [4]. Изображения третьей группы являются результатом применения всевозможных преобразований к образам первых двух. В частности, для выполнения описываемого исследования нами была синтезирована база KNI, для создания которой использовались "волновые искажения", позволяющие имитировать эффекты, возникающие при неблагоприятных окружающих условиях или недостатках аппаратуры.

Выделенные типы изображений послужили прообразами к созданию трех видов *экспертов*, под которыми в дальнейшем будем понимать классификаторы, специализирующиеся в распознавании изображений одного из типов. На основе экспертного подхода было выполнено исследование проблемы "хрупкости" по отношению к модели CNN. При этом для обучения и тестирования нейронных сетей мы использовали наиболее распространенные в области OCR базы данных.

*Базы маркированных образов цифр.* MNIST [7] – подмножество базы данных NIST, содержит изображения рукописных цифр, разделенных на тренировочную (60000 образов) и тестовую (10000) части. Изображения были получены от респондентов бюро переписи и студентов учреждений образования США, при этом в разные части были помещены образы различных авторов. Из NIST образы были смасштабированы, в прямоугольник размером 20×20, который в дальнейшем был помещен в изображение размером 32×32 пикселя, при этом должны были совпасть центр тяжести символа и геометрический центр изображения. Отметим, что на основе NIST мы также создали новую базу, названную *NIST\_HSF4*, которая содержит 55000 образов (по 11000 для каждого из десяти классов), не включенных в MNIST.

Кроме того, в исследовании применялись следующие базы:

- рукописные: USPS [8] – включает 7291 образов тренировочной и 2007 тестовой части, одноименной базы, размерности 16×16, OPTDIGITS [9] – содержит все 5620 образов одноименной базы, размерности 32×32 пикселя;

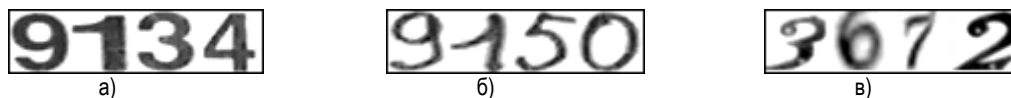
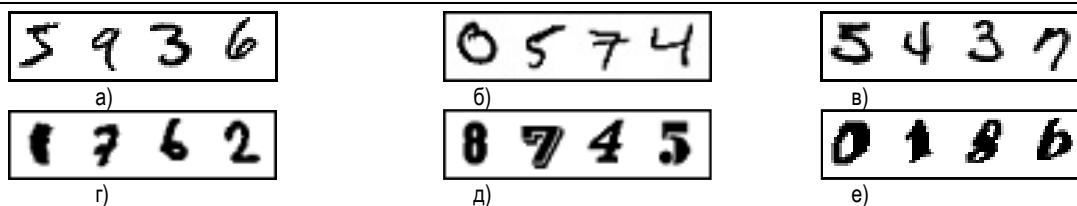


Рис. 1. Примеры изображений изолированных цифр выделенных типов: машинописных (а), рукописных (б), синтезированных (в)



MNIST (а), NIST\_HSF4 (б), USPS (в), OPTDIGITS (г), FONT (д), KNI (е)

Рис. 2. Примеры образов использованных баз данных

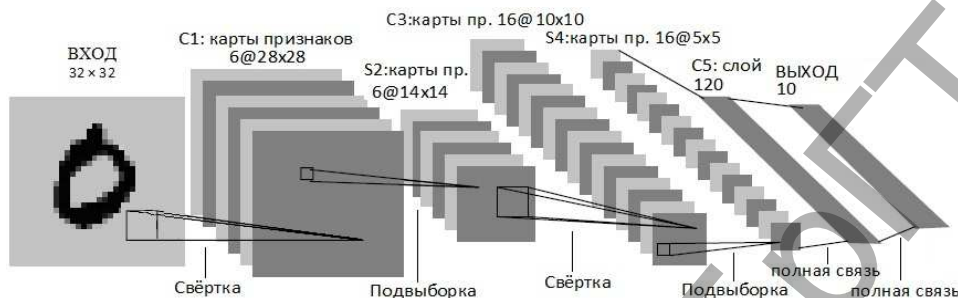


Рис. 3. Используемая сверточная нейросетевая архитектура

- машинописные объединены в одну под общим названием *FONT*, состоящую из тренировочной части, включающей 5008 шрифтовых компьютерных образов с нормальным и полужирным начертанием и 18660 образов из [10], а также тестовой части, содержащей 5008 шрифтовых курсивных образов, размерности 128×128;
- синтезированные: *KNI* – включает тренировочную (60000 образов) и тестовую (10000) части, полученные с помощью собственного программного средства на основе волновых искажений и поворотов в диапазоне ±20° из компьютерных шрифтов, с контролем попарного различия образов в каждом классе, размерности 32×32 пикселя. Примеры образов баз приведены на рисунке 2.

**Сверточные нейронные сети и проблема "хрупкости".** В качестве базовой для построения CNN экспертов нами была выбрана архитектура LeNet-5 [1], способная выделять высокоуровневые признаки образов в автоматическом режиме. Архитектура включает входной слой с 32×32 нейронами и 6 скрытых, при этом первый из них (С1) является сверточным с 6-ю картами, содержащими по 28×28 нейронов и фильтры размером 5×5. За ним следует подвыборочный слой (С2), который осуществляет усреднение откликов нейронов предыдущего по неперекрывающимся окрестностям размером 2×2 и связывает свои нейроны с 16-ю картами размера 10×10 следующего сверточного слоя (С3). Второй подвыборочный слой (С4) уменьшает размер карт признаков до 5×5. Следующий слой (С5) имеет 120 нейронов и выполняет полносвязное соединение с предыдущим. Выходной слой содержит по одному нейрону на каждый класс, т.е. 10 для образов цифр. Отметим, что описанная архитектура, изображение которой приведено на рисунке 3, является упрощенной, по сравнению с классическим LeNet-5, т.к. перед выходным в ней отсутствует RBF слой, ввиду целесообразности его применения при большом количестве классов, например для образов полного алфавита.

Обучение нейронных сетей осуществлялось с помощью модификации алгоритма обратного распространения ошибки стохастическим диагональным методом Левенберга-Марквардта, в online-режиме, с параметром обучения, равномерно уменьшающимся от 0.001 до 0.000001 в течение 68 эпох. Перед каждой из них образы тренировоч-

ного множества подвергались эластичным (с параметрами:  $\sigma = 8$ ,  $\alpha = 36$ , [2]) и аффинным искажениям (поворот на ±15°, для изображений цифр '1' и '7' – ±7°, и масштабирование в пределах ±15%, для каждой размерности отдельно). Отметим, что искажения являются важным фактором в повышении обобщающих способностей сети, при этом они предотвращают переобучение, возможное в связи с нарушением неравенства Вапника-Червоненкиса, связывающего число параметров модели и размер тренировочного множества [11].

Для построения трех CNN экспертов мы использовали рукописную (MNIST), машинописную (FONT) и синтезированную (KNI) базы данных, которые в дальнейшем будем называть главными. Имена экспертов соответствуют базам, на основе которых они были созданы, при этом их обучение проводилось на тренировочных, а проверка обобщаемости на тестовых частях баз. Результаты тестирования экспертов представлены в таблице 1.

Анализ данных таблицы приводит к следующим замечаниям: 1) использованная CNN архитектура и методика ее обучения позволили достичь более 99% точности распознавания экспертами образов своих баз; 2) максимальный уровень ошибок был показан экспертами на множестве *KNI\_train*; 3) эксперт, обученный на основе MNIST, обладал лучшей средней точностью как на тренировочных (93.88%), так и на тестовых (93.22%) частях баз.

Наконец, средняя точность экспертов (92.58%) на тестовых множествах подтвердила актуальность проблемы "хрупкости" по отношению к модели CNN.

**Комитеты как способ объединения знаний экспертов.** Имея высокоточные на своих обучающих и тестовых множествах нейронные сети с низким уровнем корректного распознавания образов других баз, логичным выглядела попытка интеграции их знаний на основе технологий коллективного распознавания. Альтернативой данному подходу являлось построение CNN на объединенном множестве обучающих частей различных баз, однако в этом случае возникает риск критичного увеличения времени обучения и/или необходимости расширения нейросетевой архитектуры.

Таблица 1. Точность распознавания (в %) CNN экспертами образов тренировочных (train) и тестовых (test) частей главных баз

CNN эксперт	MNIST_train	FONT_train	KNI_train	Среднее для 3-х	MNIST_test	FONT_test	KNI_test	Среднее для 3-х
CNN_MNIST	99.68	96.75	85.21	93.88	99.39	95.23	85.06	93.22
CNN_FONT	86.68	99.88	88.34	91.63	87.23	99.58	88.06	91.62
CNN_KNI	86.48	94.54	99.38	93.46	86.52	92.97	99.29	92.92
Среднее	90.94	97.05	90.97	92.99	91.04	95.92	90.80	92.58

Под коллективным распознаванием будем понимать задачу использования множества (комитета, ансамбля и др.) классификаторов, каждый из которых принимает решение о классе одной и той же сущности с последующим согласованием их решений с помощью некоторого алгоритма. Важным условием эффективного формирования комитета является соблюдение необходимого баланса между *точностью* и *разнообразием* членов. При этом если их точность можно оценить на валидационной выборке пространства образов, то измерение уровня разнообразия является не столь очевидным, ввиду большого числа вариантов формализации данного понятия в терминах мер [12]. Зачастую под разнообразием понимают степень некоррелированности ошибок членов комитета, существенное влияние которой подтверждается, в том числе, экспериментально. В частности Zhang продемонстрировал преимущество объединения трех классификаторов, каждый из которых обладал точностью на уровне 67% и низким уровнем коррелированности ошибок перед аналогичным объединением с точностью членов в районе 95% [13].

Немаловажным фактором эффективности комитета является схема объединения голосов членов. Существуют различные схемы голосования [12], выбор которых зависит от пространства признаков, моделей классификаторов и др. В данном исследовании были применены наиболее универсальные, для которых победителем является класс:

- 1) *максимальное* – с максимальным откликом членов комитета;
- 2) *усредняющее* – с наибольшим средним откликом членов комитета;
- 3) *большинством* – с наибольшим числом голосов членов комитета.

Ввиду наличия хороших предпосылок для формирования комитетов (единая нейросетевая архитектура, методика обучения), обеспечивающих нормировку откликов экспертов в единый числовой диапазон, на основе ранее созданных CNN нами были построены три комитета, названия которых соответствуют используемой схеме голосования членов: MAX\_COM, AVER\_COM, MAJOR\_COM. Результаты их тестирования на главных и контрольных базах представлены в таблице 2. Отметим, что для сравнения эффективности в таблице также приведены результаты системы KADMOS, являющейся коммерческим средством, специализирующимся в решении OCR задач [14]. Свободный в распространении KADWOS MINI SDKs содержит классификатор образов цифр с машинописным и рукописным начертанием ('numbers\_us.rec'), точность которого отражена в таблице.

Как можно заметить, MAX\_COM является наиболее эффективным из созданных комитетов в распознавании образов как главных, так и контрольных баз. Анализ его результатов приводит к следующим выводам: 1) точность комитета на тестовых частях главных баз превышает среднюю точность экспертов на 5.89%, что доказывает эффективность объединения их знаний; 2) средняя точность комите-

та на контрольных базах (96.50%) показывает его способность к обобщению и стабильность характеристик; 3) средняя точность комитета (97.48%) на множествах всех баз превышает аналогичную для системы KADMOS на 3.21%, что подтверждает перспективность описываемого подхода.

Таким образом, можно утверждать, что интеграция знаний CNN экспертов в комитете, позволяет существенно повысить точность распознавания. Для сравнения, Seewald в [4] с помощью объединения не экспертов, а тренировочных множеств трех рукописных баз, получил на 2.73% меньшую среднюю точность, чем лучший из созданных комитетов на множествах с различным стилем начертания. Однако следует отметить, что данный результат все еще не достаточен для универсальности сферы применения классификатора.

**Комитеты как средство повышения перекрестной точности экспертов.** Объединение знаний сверточных нейронных сетей в комитете доказало свою эффективность, продемонстрировав перспективность применения технологий коллективного распознавания в преодолении проблемы "хрупкости". В этой связи было решено изучить возможность повышения перекрестной точности экспертов (корректности распознавания образов баз, не участвующих в их обучении), основываясь на данных технологиях. А именно, оценить эффективность перехода от формы экспертов в виде единичных CNN к их комитетам и объединения частных комитетов в одном. В результате, подвергаемая анализу итоговая классификационная модель может быть названа "комитетом CNN комитетов".

Одно из возможных направлений продолжения исследования основывалось на разделении тренировочного множества на подвыборки с целью построения классификаторов на каждой из них (например, bagging, boosting и др. [12]). Однако, на наш взгляд, наиболее перспективным, учитывая особенности решаемой задачи, является не разбиение, а расширение тренировочного множества с помощью процедуры "регулярного масштабирования": выберем диапазон изменения высоты ( $h$ ) и ширины ( $w$ ) символа (например, [16, 24] для  $h$  и [10, 18] для  $w$  с шагом в 2 пикселя) и выполним масштабирование к каждому из допустимых размеров (всего возможно 25 комбинаций с фиксацией  $h$  и  $w$ , 10 – с фиксацией одного из них). На основе полученных таким образом множеств можно построить 35 CNN и объединить их в комитет (схема его использования приведена на рис. 4).

Преимуществами предложенного способа расширения тренировочного множества являются: 1) отсутствие необходимости привлекать дополнительные базы маркированных образов; 2) обеспечение высокого уровня их уникальности; 3) соответствие естественной природе различия образов (размера шрифтов, наклона и др.).

Таблица 2. Точность распознавания (в %) комитетами экспертов образов контрольных баз и тестовых частей (test) главных

Комитет CNN	MNIST_test	FONT_test	KNI_test	Среднее для 3-х	NIST_HSF 4	OPTDIGITS	USPS	Среднее для 3-х
MAX_COM	98.23	98.62	98.56	98.47	97.19	94.51	97.80	96.50
AVER_COM	97.64	99.11	98.10	98.28	96.37	93.73	96.52	95.54
MAJOR_COM	93.82	98.50	98.86	97.06	92.01	87.68	92.95	90.88
KADMOS	95.84	98.68	84.95	93.15	94.45	94.66	97.09	95.4

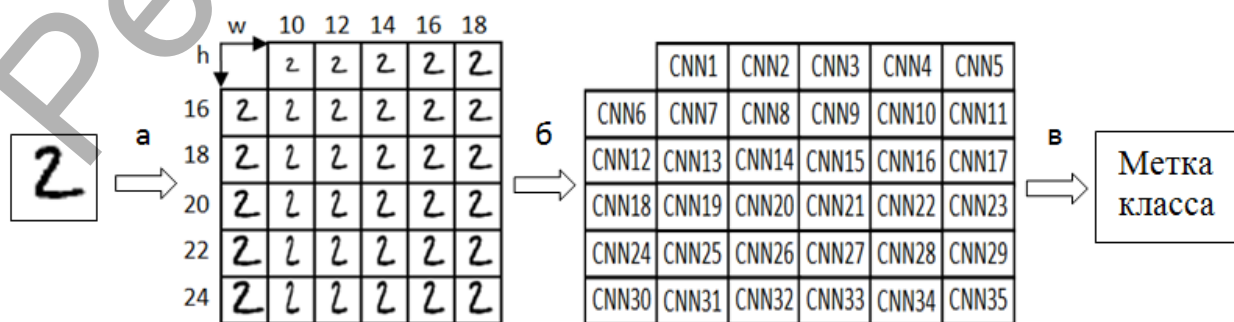


Рис. 4. Схема применения эксперта как комитета сверточных нейронных сетей: а) регулярное масштабирование; б) распознавание образов; в) объединение решений

Таблица 3. Точность распознавания (в %) частными экспертными комитетами образов тестовых частей главных баз

Частный экспертный комитет	Усредняющее голосование			Максимальное голосование			Голосование большинством		
	Mnist_test	FONT_test	KNI_test	Mnist_test	FONT_test	KNI_test	Mnist_test	FONT_test	KNI_test
10_CNN_MNIST	99.58	96.85	89.77	99.51	96.16	89.00	99.57	96.59	89.6
10_CNN_FONT	90.82	99.82	90.44	90.57	99.82	90.37	90.66	99.80	90.2
10_CNN_KNI	90.22	96.31	99.86	89.49	96.14	99.80	89.90	92.27	99.8

Таблица 4. Точность (в %) распознавания классификаторами образов контрольных баз и тестовых частей главных

Классификатор	MNIST_test	FONT_test	KNI_test	NIST_hsf4	OPTDIGITS	USPS
COM_COM	98.95	99.27	99.06	98.25	96.44	98.39
MAX_COM	98.23	98.62	98.56	97.19	94.51	97.80
KADMOS	95.84	98.68	84.95	94.45	94.66	95.4

Используя предложенную процедуру, для каждой тренировочной части трех главных баз образов (MNIST, FONT, KNI) нами были получены по 35 обучающих подмножеств. Тестовые же подмножества были сформированы аналогичным образом из тестовых частей баз и приведены к размерам, соответствующим обучающим. Таким образом, были подготовлены данные, необходимые для построения CNN. При этом отметим, что, ввиду длительности процесса обучения одной CNN (от 2 до 5 часов на ЭВМ со стандартной аппаратной конфигурацией), из требуемых 105 к моменту написания данной статьи были построены 30 сетей (по 10 для каждой базы), на основе которых и были сформированы частные комитеты экспертов, результаты тестирования которых приведены в таблице 3.

Данные в таблице показывают: 1) наиболее эффективной для формирования экспертных комитетов является усредняющая схема голосования; 2) повышение перекрестной точности комитетов с данной схемой, по сравнению с единичными CNN, в среднем составило: 2.17% для 10\_CNN\_MNIST, 2.07% для 10\_CNN\_FONT, 2.53% для 10\_CNN\_KNI; 3) повышение корректности распознавания образов главных баз в среднем составило: 2.49% для MNIST, 1.73% для FONT, 2.55% для KNI. Таким образом, предложенная методика повышения качества экспертов доказала свою эффективность. Для оценки ее вклада в решение главной задачи – преодоление проблемы "хрупкости" – три частных комитета, использующих усредняющую схему, были объединены в один (с максимальной схемой), названный COM\_COM, результаты тестирования которого приведены в таблице 4.

Основные выводы: 1) COM\_COM позволил повысить точность распознавания всех множеств (по сравнению с MAX\_COM) в среднем для главных баз на 0.62%, для контрольных – на 1.19%, что доказывает значимость перекрестной точности экспертов в преодолении проблемы "хрупкости"; 2) средняя точность COM\_COM на всех базах (98.39%) превышает аналогичную для MAX\_COM (97.48%) на 0.91%, что показывает гибкость предложенного подхода, возможность его развития; 3) невысокий результат в распознавании базы OPTDIGITS подтверждает необходимость разработки более совершенной единой методики предобработки образов, инвариантной, в частности, к поворотам и толщине; 4) средняя точность COM\_COM превышает аналогичную для KADMOS (94.27%) на 4.12%, что говорит о достижении им более высокого уровня универсальности по сравнению с коммерческими OCR системами; 5) продолжение исследований (в частности, создание полного множества CNN) позволяет надеяться на дальнейшее увеличение эффективности классификатора и перспективность его промышленного приложения.

**Селекция членов как значимый фактор эффективности комитетов.** Идея повышения перекрестной точности экспертов на основе "регулярного масштабирования" была реализована не в полной мере (из запланированных 35 сетей для каждого использовалось 10), однако даже в своей сокращенной форме она позволила добиться существенного прогресса в решении главной задачи. Данный факт, наряду с вычислительной сложностью процесса обучения и применения большого числа CNN, а также схожесть характеристик

членов в рамках одного экспертного комитета, привели к выводу: увеличение количества членов комитетов должно сопровождаться процедурой их селекции.

Исследование данного вопроса было решено выполнить с помощью базы MNIST, процедуры регулярного масштабирования и усредняющей схемы учета голосов. В результате были построены 10 подмножеств тренировочной части MNIST, обладающие следующей вариацией размеров объемлющих прямоугольников образов: ( $h = 16, 18, 20, 22, 24, w =$  пропорционально  $h$ ), ( $h = 20, w = 10, 12, 14, 16, 18$ ). Используя данные подмножества, были обучены 10 CNN, точность распознавания тестовой части MNIST которыми составила: 99.14%, 99.39%, 99.39%, 99.35%, 99.36%, 99.28%, 99.33%, 99.41%, 99.34%, 99.44%.

Целью селекции является выбор подкомитета, обладающего более высокой или, по крайней мере, равной точностью, чем комитет. Качество селективных алгоритмов зависит от оценки вклада членов комитета в его эффективность, при этом основными анализируемыми характеристиками являются точность и разнообразие. Так авторы работы [13] пришли к заключению, что члены, обладающие большим (меньшим) количеством правильных (ошибочных) предсказаний в меньшинстве, наиболее (наименее) полезны (вредны) при формировании комитета. Предложенный ими селективный алгоритм, названный EPIC, обладает рядом преимуществ по сравнению с другими, поэтому он и был выбран для выполнения экспериментальной работы. Его выходом является перечень членов комитета в порядке убывания их вкладов в общую эффективность.

Нами были опробованы два способа интерпретации выхода: 1) путем пошагового добавления членов в порядке перечня формируются 10 подкомитетов и выбирается самый точный из них; 2) подкомитет инициализируется первым членом перечня, добавление остальных ведется до тех пор, пока может быть найден член, увеличивающий точность подкомитета (перечень просматривается сверху, член может быть добавлен в подкомитет единожды). В результате применения алгоритма EPIC были сформированы два подкомитета, CNN в которых обучались на образах следующих размеров: первым способом – ( $h = 16, 18, 24, w =$  пропорционально), ( $h = 20, w = 10, 14$ ), вторым – аналогичные + ( $h = 20, w =$  пропорционально). Точность распознавания ими тестового MNIST составила 99.59% и 99.6%, что незначительно превышало данный показатель исходного комитета (99.58%), однако количество членов по сравнению с ним было уменьшено в 2 раза.

В алгоритме EPIC существенную роль играет корректность распознавания членами тестовых образов, причем она измеряется в бинарной форме. Возникло предположение о целесообразности использования вместо бинарных, нормированных от  $[-1, 1]$  к  $[0, 1]$  значений выходного вектора CNN. Так, модифицированная версия алгоритма EPIC и два описанных способа интерпретации выхода привели к формированию одинаковых подкомитетов, члены которых обучались на образах следующих размеров: ( $h = 16, 24, w =$  пропорционально), ( $h = 20, w = 10, 18$ ). Хотя средняя индивидуальная точность членов составила 99.30%, их комитет позволил получить 99.64% корректного распознавания тестового MNIST! Причиной



достижения данного результата является способность модифицированного алгоритма качественно оценивать вклад членов и поддерживать баланс между их точностью и разнообразием. Так с позиций точности им были выбраны 10-я, 1-я, 9-я и 6-я сети (перечислены в порядке уменьшения вкладов), однако со стороны разнообразия данный выбор не случаен, т.к. сети обучались на символах с максимально допустимыми вариациями высоты [16, 24] и ширины [10, 18] для данного эксперимента.

Подчеркнем, что полученная точность распознавания тестового MNIST достаточно уникальна, т.к. она является наивысшей для классификаторов архитектуры LeNet-5 и других не нейросетевых, включая SVM, K-NN и др., при этом международное соперничество по распознаванию образов данной базы, проводимое с 1998 г. по настоящее время десятками ученых, является одним из наиболее престижных в области машинного обучения. В сравнении с полученным, более высоким является лишь результат швейцарской команды исследователей (0.35%, 0.27%, и 0.23%, 2010–2012 гг. [3]), который был получен благодаря применению более громоздких архитектур, например, MLP с числом нейронов по слоям: 1-20-40-60-80-100-120-120-10. Кроме того, ими было использовано мощное оборудование, в частности несколько расчетных GPU, что позволило повысить скорость обучения CNN в десятки раз. Отметим критичность данного фактора в проведении эффективной экспериментальной работы, связанной с созданием нейросетевых классификаторов.

Таким образом, можно сделать следующие выводы: 1) селективные алгоритмы способны уменьшить количество членов комитета, а следовательно, и вычислительную сложность; 2) обеспечивая хороший баланс между точностью и разнообразием членов, они способны сформировать более эффективное объединение, чем исходное; 3) качество работы селективных алгоритмов может быть повышено при учете особенностей модели классификаторов. На основании указанных выводов можно предположить, что данные алгоритмы могут оказаться существенным фактором в преодолении проблемы "хрупкости".

**Заключение.** Основные результаты исследования: 1) обнаружена перспективная трактовка задачи построения универсального классификатора образов цифр как преодоление проблемы "хрупкости", учет которой может способствовать прогрессу в области построения промышленных систем распознавания; 2) проведена теоретическая и экспериментальная работа, показавшая эффективность применения комитетов CNN, обученных на базах с различным стилем начертания

образов в сочетании с регулярным варьированием их ширины и высоты, для разрешения данной проблемы; 3) сформирован комитет CNN обладающий уникальной точностью распознавания тестового MNIST (0.36%); 4) построен классификатор, средняя точность которого (98.39%) соответствует уровню коммерческих OCR.

#### СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Головкин, В.А. Нейронные сети: обучение, организация и применение. – М., 2001.
2. LeCun, Y. Gradient-Based Learning Applied to Document Recognition // Proceedings of the IEEE, 86(11), 1998 / Y. LeCun, L. Bottou, Y. Bengio and P. Haffner – P. 2278–2324.
3. Simard, P. Best practices for convolutional neural networks applied to visual document analysis // ICDAAR / P. Simard, D. Steinkraus, and J. Platt, 2003. – P. 958–963.
4. Dan Ciresan, Ueli Meier, Jürgen Schmidhuber Multi-column deep neural networks for image classification // CVPR, 2012. – P. 3642–3649.
5. Alexander, K. Seewald On the Brittleness of Handwritten Digit Recognition Models // ISRN Machine Vision, 2012.
6. Palacios, R. A system for processing handwritten bank checks automatically // Image and Vision Computing, 2008 / R. Palacios, A. Gupta – P. 1297–1313.
7. MNIST database. – Режим доступа: <http://yann.lecun.com/exdb/mnist/index.html>.
8. Hastie, T. The Elements of Statistical Learning. Data Mining, Inference and Prediction / T. Hastie, R. Tibshirani, J. Friedman – New York: Springer, 2001.
9. Optdigits database. – Режим доступа: <http://mlearn.ics.uci.edu/databases/optdigits/>.
10. Weinman, J.J. Scene text recognition using similarity and a lexicon with sparse belief propagation / J.J. Weinman, E. Learned-Miller, A. Hanson // IEEE on PAMI, 2009. – P. 1733–1746.
11. Bishop, C.M. Neural Networks for Pattern Recognition. – Oxford University, 1995.
12. Kuncheva L.I. Combining Pattern Classifiers. Methods and Algorithms. – Wiley, 2004.
13. Zhenyu, Lu Ensemble pruning via individual contribution ordering // KDD, 2010 / Lu Zhenyu, Wu Xindong, Zhu Xingquan, Josh Bongard – P. 871–880.
14. KADMOS recognition software. Режим доступа: <http://www.rerecognition.com>.

Материал поступил в редакцию 05.10.12

#### KUZMITSKY N.N. Actual questions of use of convolutional neural networks and their committees in recognition of digit patterns

The task of creation of the universal qualifier of digit patterns on a basis of convolutional neural networks is investigated. The analysis of "brittleness" of models of systems of statistical training as main problem in the solution of the specified task is made. Prospects of use of committees as integration tool of knowledge of neural networks and increase of their accuracy is shown. Efficiency of application of bases with different writing style of patterns in combination to regular variation of their width and heights for overcoming of a problem of "brittleness" is proved. The qualifier with average accuracy of recognition of digit patterns over 98 % is created. The committee of the neural networks, allowing to receive 0.36 % of mistakes on a test part of the MNIST base is created.

УДК 004.8.032.26

**Кочурко П.А., Головкин В.А.**

### АНСАМБЛЬ НЕЙРОСЕТЕВЫХ ДЕТЕКТОРОВ В СИСТЕМАХ ОБНАРУЖЕНИЯ АТАК

**Введение.** Среди задач, которые решают системы обнаружения атак, основной является обнаружение сетевой активности, которая может нанести урон информационной безопасности вычислительной системы. Подобная активность, выражающаяся в сетевом трафике, отличном от нормального, может считаться атакой. СОА производит перехват трафика, обрабатывает его с формированием характеристик, значимых для последующего анализа. Полученные характеристики трафика проходят предварительную обработку, после чего подаются

на вход ИНС, производящих обнаружение атак технологиями обнаружения аномалий и злоупотреблений. Подобная схема позволяет СОА обнаруживать новые, неизвестные ранее атаки.

Системы обнаружения атак реализуют технологии обнаружения аномалий и обнаружения некорректного поведения или злоупотреблений. Применяемые большинством современных систем методы – на основе правил, статистические и другие – недостаточно эффективно обнаруживают атаки, особенно модифицированные и неиз-

**Кочурко Павел Анатольевич, к.т.н., доцент кафедры интеллектуальных информационных технологий Брестского государственного технического университета.**

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.