

**МОБИЛЬНОЕ РАСПОЗНАВАНИЕ НОМИНАЛОВ РОССИЙСКИХ БАНКНОТ
ДЛЯ СЛЕПЫХ ЛЮДЕЙ**Д.А. Суворов¹Р.А. Жуков¹Д.О. Тетерюков¹М.В. Мозговой²А.В. Волков²

dmitry.suvorov@skolkovotech.ru

roman.zhukov@skolkovotech.ru

D.Tsetserukou@skoltech.ru

mozgovoy@bmstu.ru

avv@bmstu.ru

¹ Сколковский институт науки и технологий, Сколково,
Москва, Российская Федерация

² МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Представлена система классификации российских банкнот для слепых людей по фото на основе методов глубокого обучения с использованием техники *knowledge transfer*. Проведен сравнительный анализ производительности и точности подходов с использованием архитектур *ResNet-50*, *VGG-19* и *Inception-v3* для первичного извлечения признаков с фото купюр. Описан процесс создания базы для обучения системы. Проведено тестирование системы на настольном и мобильном процессорах. Наилучшую точность распознавания показала система, построенная на базе архитектуры для извлечения признаков *ResNet-50*. По производительности система на базе *ResNet-50* оказалась на втором месте, незначительно уступив системе на базе *Inception-v3*, которая при этом показала очень низкую точность классификации (78 %). Точность классификации решения на базе *ResNet-50* составила 99,5 %, что является приемлемым для ее использования в реальных условиях

Ключевые слова

Распознавание, банкнота, глубокое обучение, передача знаний, машинное обучение, распознавание образов

Поступила в редакцию 20.07.2017

© МГТУ им. Н.Э. Баумана, 2018

Введение. По данным Всемирной организации здравоохранения, в 2010 г. в мире насчитывалось 39 млн слепых людей [1]. Одна из проблем, с которой сталкиваются такие люди в повседневной жизни, — идентификация номиналов банкнот, которые они получают при совершении платежных операций. Вследствие одинаковых размеров и текстуры их тяжело идентифицировать на ощупь. В настоящее время известны методы идентификации банкнот с применением специализированного сканирующего оборудования [2, 3] (устанавливается в купюроприемниках, банкоматах или в счетных машинах), а также методы, использующие данные обычной видеокамеры [4–8]. Специализированное сканирующее оборудование не подходит для повседневного использования слепыми людьми ввиду высокой стоимости и низкой мобильности. Методы на основе анализа видеоданных долж-

ны обладать высокой производительностью, чтобы они могли работать на мобильных процессорах архитектуры *arm7*, которая на сегодняшний день наиболее популярна у производителей смартфонов и планшетных компьютеров.

Ранее при анализе видео для распознавания номиналов банкнот использовали методы на основе сегментации и корреляции с образцом [4], методы *SURF* [5, 6], извлечения специализированных признаков и их классификации с помощью нейронной сети [7], сегментации и вычисления евклидовых расстояний в *HSV*-пространстве [8]. Перечисленные методы обладают высокой чувствительностью к внешним условиям и параметрам камеры, плохо воспринимают сильно мятые или рваные банкноты.

Цель настоящей работы — создание системы распознавания российских банкнот для слепых людей, которая может функционировать на смартфоне. Предложен метод классификации банкнот на основе глубокого машинного обучения с использованием техники *knowledge transfer* [9]. Результаты его тестирования показали очень высокую устойчивость к изменению внешних условий, смятию или повреждению банкнот. Кроме того, метод обладает высокой производительностью, достаточной для работы на мобильных процессорах, что подтверждено его тестированием на смартфоне с операционной системой (ОС) *Android*. Описан массив данных, созданный для обучения системы. Приведен процесс обучения и анализ трех архитектур для решения поставленной задачи. Приведен результат разработки и тестирования приложений под ОС *Linux* и *Android* на базе выбранной архитектуры. Сделаны основные выводы по работе.

Данные для обучения. Для решения задачи была сформирована небольшая база фото российских банкнот (рис. 1), состоящая из 903 фотографий. Фотографии банкнот были выполнены при различном освещении и положении банкнот, на разном фоне, с частичным перекрытием банкнот, намеренно плохого качества. Это сделано для того, чтобы, обучившись на массивах, система могла наиболее качественно распознавать банкноты в реальных условиях. Поскольку для использования методов глубокого обучения необходимо наличие больших объемов обучающих данных, их объем был увеличен в 8 раз с помощью геометрических преобразований. Каждая банкнота была повернута 3 раза на 90° , затем было сделано их зеркальное отображение и повернуто на 90° еще 3 раза.

Обучение классификатора. Итоговая база, состоящая из 7 224 фото, оставалась довольно небольшой по меркам методов глубокого обучения, поэтому была использована техника *knowledge transfer*, т. е. были взяты глубокие нейронные сети,

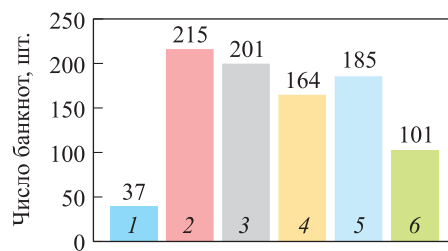


Рис. 1. Состав базы данных банкнот номиналом 10 (1), 50 (2), 100 (3), 500 (4), 1000 (5) и 5000 (6) руб.

обученные для классификации на большом объеме данных большого числа классов. Их последние слои были изменены в соответствии с требованиями задачи и обучены на собранной базе без изменения весов начальных слоев. Эксперименты проводились на языке *Python* с использованием библиотек *Theano* и *Lasagne*.

Применены архитектуры *ResNet-50* [10], *VGG-19* [11] и *Inception-v3* [12], обученные на базе *ImageNet* [13]. Последний полносвязный слой был удален из каждой базовой нейронной сети и заменен структурой (рис. 2), в которую входит один скрытый полносвязный слой с 1000 нейронами, *Dropout* и нелинейностью *ReLU* для формирования финального вектора признаков, описывающего входное изображение, и выходной слой с шестью элементами и нелинейностью *SoftMax* для классификации изображений. Размер входного слоя каждой базовой архитектуры был изменен на 320×240 пикселей, что соответствует разрешению изображения, которое регистрируется видеокамерой.



Рис. 2. Структура, добавленная вместо выходного слоя в каждую из базовых архитектур нейронных сетей

Обучение проводилось с помощью алгоритма оптимизации *Adam* [14] с низким значением скорости обучения, которое в комбинации с *Dropout* позволяло долгое время не наблюдать переобучения. В качестве функции потерь использована кросс-энтропия. Результат обучения системы на базе *ResNet-50* приведен на рис. 3. Система показала точность 99,5 % и оказалась лучшей по качеству.

Дополнительно для контроля качества классификации была выполнена визуализация векторов признаков с помощью метода *t-SNE* [15]. На вход нейронной сети подавали все изображения из базы данных. Размерность векторов выходов предпоследнего слоя была снижена с 1 000 до 32 с помощью метода главных компонент. Для новых 32-мерных векторов выполнено преобразование *t-SNE* для снижения их размерности до 2. Кластеры данных хорошо разделены друг от друга, что свидетельствует о высоком качестве классификатора (рис. 4).

Аналогичная система реализована на базе архитектуры *VGG-19*, предварительно обученной на *ImageNet*. Система дает более низкую точность, чем система на базе архитектуры *ResNet-50* — около 95 %. Кроме того, эта архитектура обучалась в несколько раз медленнее, чем архитектура *ResNet-50* на том же аппаратном обеспечении. Результат ее обучения показан на рис. 5.

Для анализа результатов обучения системы на базе архитектуры *VGG-19* также была проведена визуализация векторов признаков с помощью метода *t-SNE*. В нижней части визуализации произошло смешение образцов из разных кластеров (рис. 6), в результате чего имеет место более низкая точность, чем у *ResNet-50*.

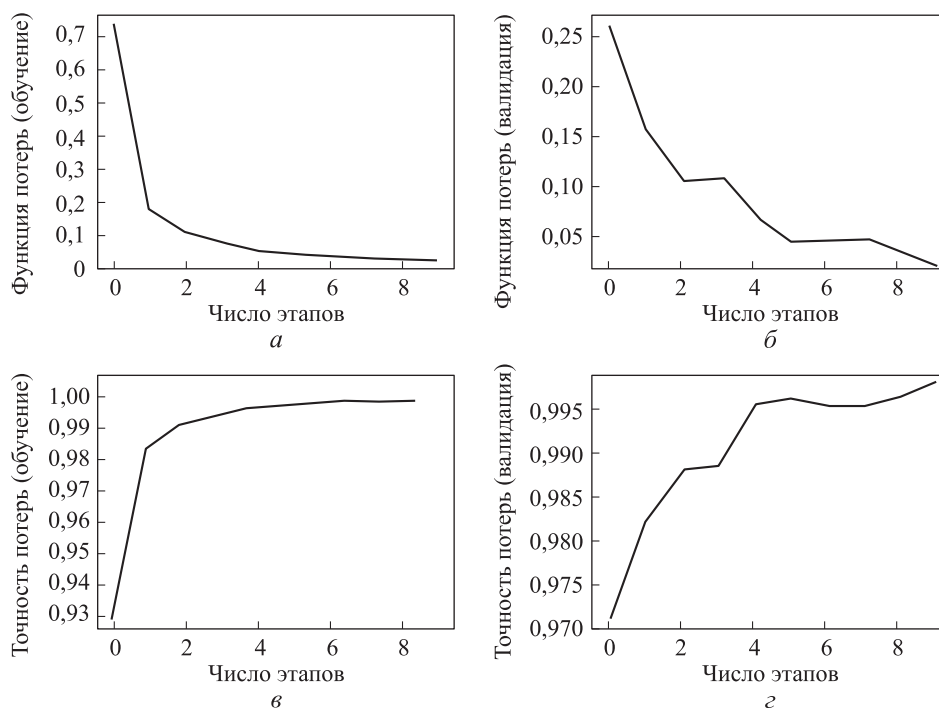


Рис. 3. Результат обучения системы на базе архитектуры *ResNet-50*:

a, б — функции потерь на обучающих и тестовых данных; *в, з* — точность классификации обучающих и тестовых данных

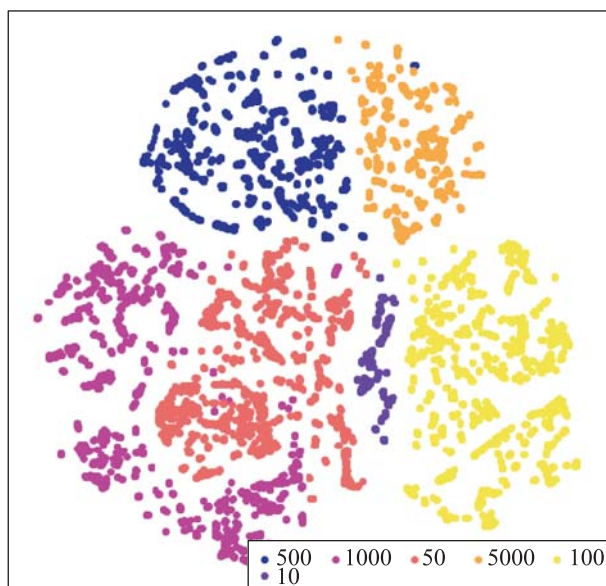


Рис. 4. Визуализация методом *t-SNE* векторов признаков на предпоследнем слое системы на базе архитектуры *ResNet-50* для банкнот номиналом 10, 50, 100, 500, 1000 и 5000 руб.

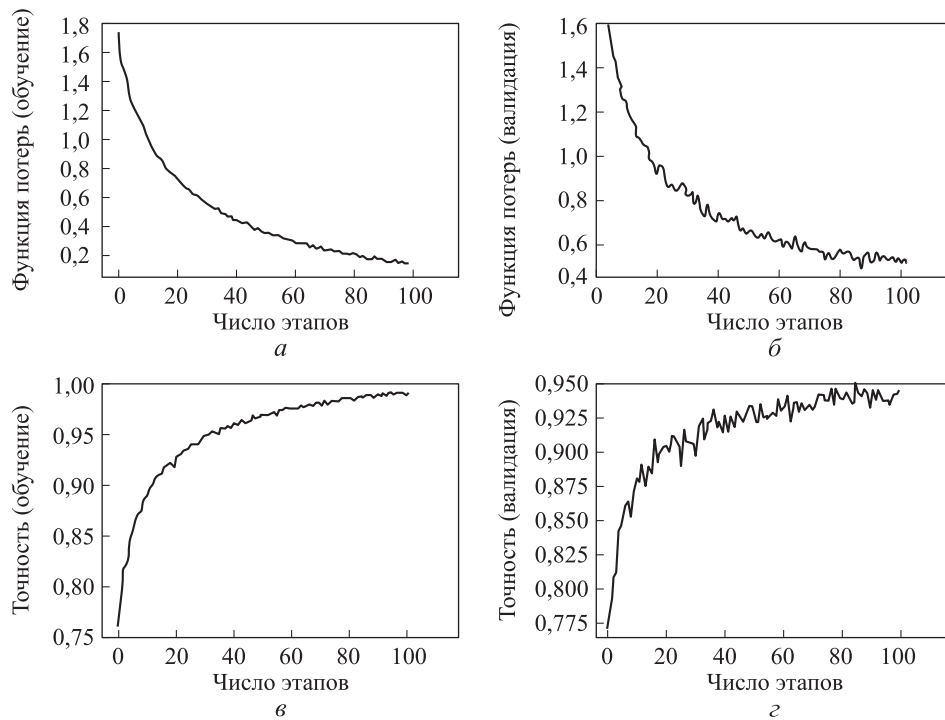


Рис. 5. Результат обучения системы на базе архитектуры VGG-19:

a, б — функции потерь на обучающих и тестовых данных; *в, г* — точность классификации обучающих и тестовых данных

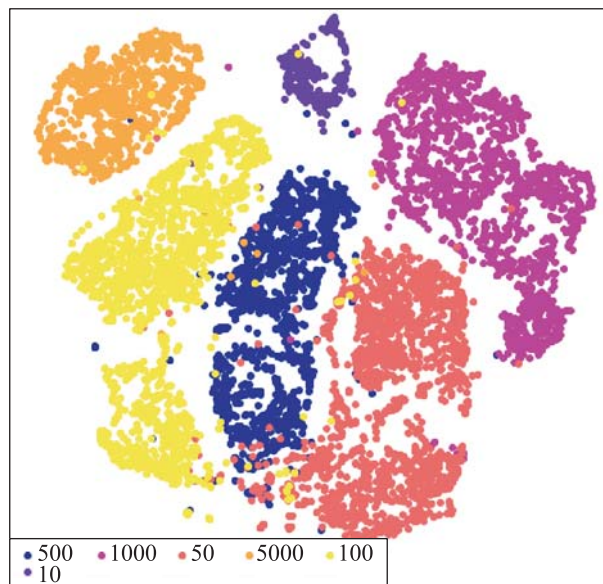


Рис. 6. Визуализация методом *t-SNE* векторов признаков на предпоследнем слое системы на базе архитектуры VGG-19 для банкнот номиналом 10, 50, 100, 500, 1000 и 5000 руб.

Архитектура на базе *Inception-v3* показала неудовлетворительное качество работы в выбранной конструкции. Точность классификации — около 78 %. Результат обучения показан на рис. 7.

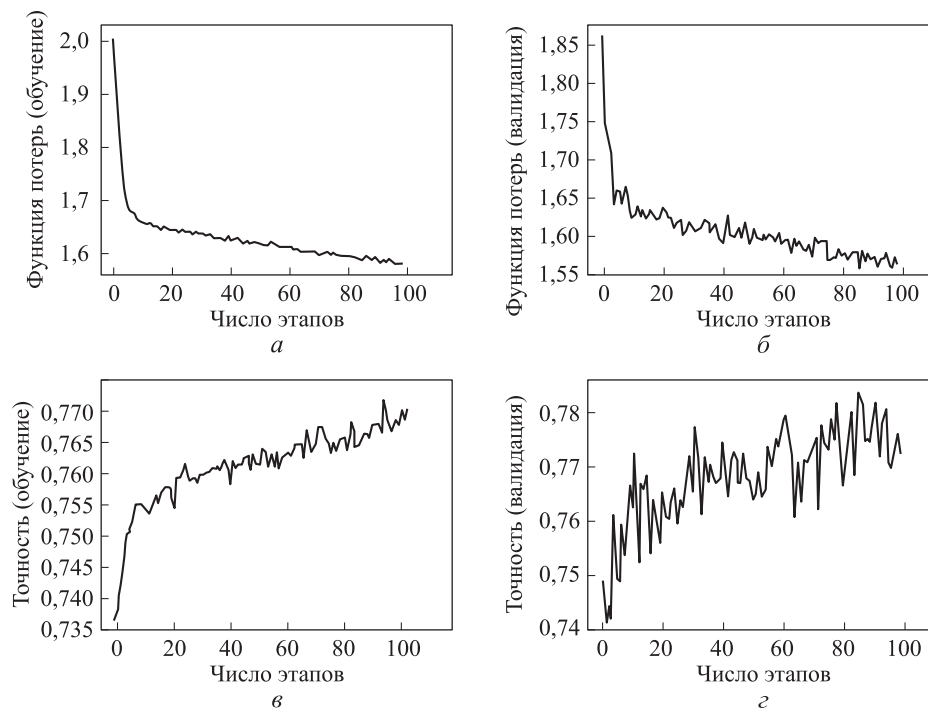


Рис. 7. Результат обучения системы на базе архитектуры *Inception-v3*:

а, б — функции потерь на обучающих и тестовых данных; в, г — точность классификации обучающих и тестовых данных

При визуализации векторов признаков методом *t-SNE* отсутствует четкая кластерная структура (рис. 8), что свидетельствует о непригодности рассмотренной системы для классификации банкнот.

При сравнении результатов обучения рассмотренных систем ясно, что система на базе архитектуры *ResNet-50* для первичного выделения признаков на изображении не только обладает наибольшей точностью, но и обучается за наименьшее число этапов, что значительно сокращает требуемое вычислительное время в процессе разработки.

Классификация в реальном времени. Дополнительно была проверена скорость классификации одного кадра на 64-разрядном процессоре. Решение на базе архитектуры *ResNet-50* показало приемлемый результат за 1,3 с на один кадр (рис. 9).

Для проверки работоспособности в реальных условиях реализовано приложение для ОС *Android* на языке *Java* с использованием библиотеки *Tensorflow*. Результаты его тестирования показали, что система хорошо справляется с типичными банкнотами, редко ошибаясь только при сильном смятии банкноты или демонстрации очень маленькой ее части.

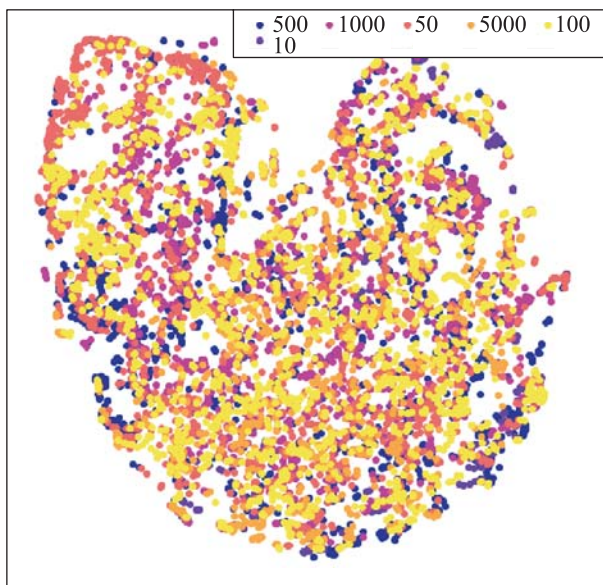


Рис. 8. Визуализация методом *t-SNE* векторов признаков на предпоследнем слое системы на базе архитектуры *Inception-v3* для банкнот номиналом 10, 50, 100, 500, 1000 и 5000 руб.

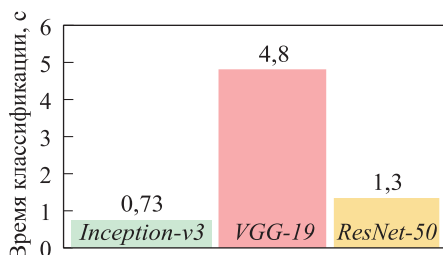


Рис. 9. Сравнение времени классификации одного кадра на базе трех реализованных архитектур

Заключение. Разработана и успешно протестирована система классификации банкнот. Она основана на технологии глубокого обучения и алгоритмически не привязана к конкретной валюте, однако в рамках проделанной работы все эксперименты были проведены для российских банкнот. Далее система может быть использована для любой другой мировой валюты. Показано, что для обучения системы классификации денег на небольшое число классов (использовано шесть классов банкнот) достаточно относительно небольшого обучающего массива данных вследствие использования техники *knowledge transfer*. При этом система успешно справляется с классификацией поврежденных или смятых купюр. Большую практическую ценность имеет факт, что система за приемлемое время может работать на мобильном процессоре архитектуры *arm7* под управлением ОС *Android*, что свидетельствует о возможности ее использования людьми с ограниченными возможностями здоровья по зрению в России уже в настоящее время.

ЛИТЕРАТУРА

1. *Pascolini D., Mariotti S.P.* New estimates of visual impairment and blindness: 2010 // *British Journal of Ophthalmology*. 2011. Vol. 96. No. 5.
URL: <http://www.who.int/blindness/estimates2011.pdf> (дата обращения: 15.07.2017).
2. *Bruna A., Farinella G.M., Guarnera G.C., Battiato S.* Forgery detection and value identification of euro banknotes // *Sensors*. 2013. Vol. 13. Iss. 2. P. 2515–2529. DOI: 10.3390/s130202515
3. *A high-performance banknote recognition system based on a one-dimensional visible light line sensor* / Y.H. Park, S.Y. Kwon, T.D. Pham, K.R. Park, et al. // *Sensors*. 2015. Vol. 15. Iss. 6. P. 14093–14115. DOI: 10.3390/s150614093
4. *Semary N.A., Fadl S.M., Essa M.S., Gad A.F.* Currency recognition system for visually impaired: Egyptian banknote as a study case // *Proc. Int. Conf. on Information and Communication Technology and Accessibility*. DOI: 10.1109/ICTA.2015.7426896
5. *Hasanuzzaman F.M., Yang X., Tian Y.* Robust and effective component-based banknote recognition by SURF features // *WOCC*. 2011. DOI: 10.1109/WOCC.2011.5872294
6. *Singh S., Choudhury S., Vishal K., Jawahar C.V.* Currency recognition on mobile phones // *Proc. 22nd Int. Conf. on Pattern Recognition*. 2014. P. 2661–2666.
URL: http://web2py.iit.ac.in/research_centres/publications/download/inproceedings.pdf.9797adb46eb9d9a7.5375726979613230313443757272656e63792e706466.pdf (дата обращения: 15.07.2017).
7. *Parlouar R., Dramas F., Macé M.J.-M., Jouffrais Ch.* Assistive device for the blind based on object recognition: An application to identify currency bills // *Proc. 11th Int. ACM SIGACCESS Conf. on Computers and Accessibility*. 2009. P. 227–228.
URL: https://www.irit.fr/~Marc.Mace/pdfs/parlouar_r_09_227.pdf (дата обращения: 15.07.2017).
8. *Bhurke C., Sirdeshmukh M., Kanitkar M.S.* Currency recognition using image processing // *Int. J. of Innovative Research in Computer and Communication Engineering*. 2015. Vol. 3. No. 5. P. 4418–4422.
9. *Gutstein S., Fuentes O., Freudenthal E.* Knowledge transfer in deep convolutional neural nets // *Proc. of Twentieth Int. Florida Artificial Intelligence Research Society Conf.* 2007. P. 104–109.
URL: <http://www.cs.utep.edu/ofuentes/FLAIRS07GutsteinS.pdf> (дата обращения: 15.07.2017).
10. *He K., Zhang X., Ren S., Sun J.* Deep residual learning for image recognition // *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*. 2016. DOI: 10.1109/CVPR.2016.90
11. *Simonyan K., Zisserman A.* Very deep convolutional networks for large-scale image recognition // *Proc. 3rd Int. Conf. on Learning Representations*. 2015.
URL: <https://arxiv.org/pdf/1409.1556.pdf> (дата обращения: 15.07.2017).
12. *Rethinking the inception architecture for computer vision* / C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna // *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*. 2016. P. 2818–2826. DOI: 10.1109/CVPR.2016.308
13. *ImageNet large scale visual recognition challenge* / O. Russakovsky, J. Deng, H. Su, J. Krause, et al. // *Int. J. of Computer Vision*. 2015. Vol. 115. Iss. 3. P. 211–252.
DOI: 10.1007/s11263-015-0816-y
14. *Kingma D.P., Ba J.* Adam: A method for stochastic optimization // *Proc. Int. Conf. on Learning Representations ICLR*. 2015. URL: <https://arxiv.org/pdf/1412.6980.pdf> (дата обращения: 15.07.2017).

15. *Maaten L., Hinton G.* Visualizing data using t-SNE // *Journal of Machine Learning Research*. 2008. No. 9. P. 2579–2605.

URL: <http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>

Суворов Дмитрий Андреевич — аспирант, Сколковский институт науки и технологий (Российская Федерация, 143026, Москва, Сколково, ул. Нобеля, д. 3).

Жуков Роман Алексеевич — аспирант, Сколковский институт науки и технологий (Российская Федерация, 143026, Москва, Сколково, ул. Нобеля, д. 3).

Тетерюков Дмитрий Олегович — канд. техн. наук, старший преподаватель, Сколковский институт науки и технологий (Российская Федерация, 143026, Москва, Сколково, ул. Нобеля, д. 3).

Мозговой Михаил Владимирович — заместитель директора по методической работе Головного учебно-исследовательского и методического центра профессиональной реабилитации лиц с ограниченными возможностями здоровья (инвалидов) МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Волков Алексей Васильевич — директор Ресурсного учебно-методического центра по обучению инвалидов и лиц с ограниченными возможностями здоровья МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Просьба ссылаться на эту статью следующим образом:

Мобильное распознавание номиналов российских банкнот для слепых людей / Д.А. Суворов, Р.А. Жуков, Д.О. Тетерюков, М.В. Мозговой, А.В. Волков // *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*. 2018. № 1. С. 94–104.

DOI: 10.18698/0236-3933-2018-1-94-104

RECOGNITION OF RUSSIAN BANKNOTE NOMINAL VALUES BY MOBILE DEVICES FOR BLIND PEOPLE

D.A. Suvorov¹

R.A. Zhukov¹

D.O. Teteryukov¹

M.V. Mozgovoy²

A.V. Volkov²

dmitry.suvorov@skolkovotech.ru

roman.zhukov@skolkovotech.ru

D.Tsetserukou@skoltech.ru

mozgovoy@bmstu.ru

avv@bmstu.ru

¹Skolkovo Institute of Science and Technology, Skolkovo, Moscow, Russian Federation

²Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

This paper focuses on the system of recognition the nominal values of Russian banknotes by photos for blind people. The system uses the *knowledge transfer* technique and deep learning methods. In our research we compared and analyzed the performance and accuracy of approaches by using the *ResNet-50*, *VGG-19* and *Inception-v3* architectures for the primary feature extraction from photos. After that we developed three prototypes based on these architectures and tested the system on desktop and mobile processors. The

Keywords

Recognition, banknote, deep learning, knowledge transfer, machine learning, image recognition

system based on the *ResNet-50* architecture showed the best recognition accuracy. As for its efficiency, it appeared to be worse than that of the system based on *Inception-v3* architecture. However, the *Inception-v3* architecture showed very low accuracy of 78 %. Findings of the research show that *ResNet-50* architecture could be used in real life conditions due to the accuracy of the solution based on it

Received 20.07.2017

© BMSTU, 2018

REFERENCES

- [1] Pascolini D., Mariotti S.P. New estimates of visual impairment and blindness: 2010. *British Journal of Ophthalmology*, 2011, vol. 96, no. 5. Available at: <http://www.who.int/blindness/estimates2011.pdf> (accessed: 15.07.2017).
- [2] Bruna A., Farinella G.M., Guarnera G.C., Battiato S. Forgery detection and value identification of euro banknotes. *Sensors*, 2013, vol. 13, iss. 2, pp. 2515–2529. DOI: 10.3390/s130202515
- [3] Park Y.H., Kwon S.Y., Pham T.D., Park K.R., Jeong D.S., Yoon S.S. A high-performance banknote recognition system based on a one-dimensional visible light line sensor. *Sensors*, 2015, vol. 15, iss. 6, pp. 14093–14115. DOI: 10.3390/s150614093
- [4] Semary N.A., Fadl S.M., Essa M.S., Gad A.F. Currency recognition system for visually impaired: Egyptian banknote as a study case. *Proc. Int. Conf. on Information and Communication Technology and Accessibility*, 2015. DOI: 10.1109/ICTA.2015.7426896
- [5] Hasanuzzaman F.M., Yang X., Tian Y. Robust and effective component-based banknote recognition by SURF features. *WOCC*, 2011. DOI: 10.1109/WOCC.2011.5872294
- [6] Singh S., Choudhury S., Vishal K., Jawahar C.V. Currency recognition on mobile phones. *Proc. 22nd Int. Conf. on Pattern Recognition*, 2014, pp. 2661–2666. Available at: http://web2py.iit.ac.in/research_centres/publications/download/inproceedings.pdf.9797adb46eb9d9a7.5375726979613230313443757272656e63792e706466.pdf (accessed: 15.07.2017).
- [7] Parlouar R., Dramas F., Macé M.J-M, Jouffrais Ch. Assistive device for the blind based on object recognition: An application to identify currency bills. *Proc. 11th Int. ACM SIGACCESS Conf. on Computers and Accessibility*, 2009, pp. 227–228. Available at: https://www.irit.fr/~Marc.Mace/pdfs/parlouar_r_09_227.pdf (accessed: 15.07.2017).
- [8] Bhurke C., Sirdeshmukh M., Kanitkar M.S. Currency recognition using image processing. *Int. J. of Innovative Research in Computer and Communication Engineering*, 2015, vol. 3, no. 5, pp. 4418–4422.
- [9] Gutstein S., Fuentes O., Freudenthal E. Knowledge transfer in deep convolutional neural nets. *Proc. of Twentieth Int. Florida Artificial Intelligence Research Society Conf.*, 2007, pp. 104–109. Available at: <http://www.cs.utep.edu/ofuentes/FLAIRS07GutsteinS.pdf> (accessed: 15.07.2017).
- [10] He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2016. DOI: 10.1109/CVPR.2016.90
- [11] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proc. 3rd Int. Conf. on Learning Representations*, 2015. Available at: <https://arxiv.org/pdf/1409.1556.pdf> (accessed: 15.07.2017).
- [12] Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826. DOI: 10.1109/CVPR.2016.308

[13] Russakovsky O., Deng J., Su H., Krause J., Sanjeev S., Ma S., Zhiheng H., Karpathy A., Khosla A., Bernstein M., Berg A.C., Li F.-F. ImageNet large scale visual recognition challenge. *Int. J. of Computer Vision (IJCV)*, 2015, vol. 115, iss. 3, pp. 211–252. DOI: 10.1007/s11263-015-0816-y

[14] Kingma D.P., Ba J. Adam: A method for stochastic optimization. *Proc. Int. Conf. on Learning Representations ICLR*, 2015.

Available at: <https://arxiv.org/pdf/1412.6980.pdf> (accessed: 15.07.2017).

[15] Maaten L., Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, no. 9, pp. 2579–2605.

Available at: <http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>

Suvorov D.A. — post-graduate student, Skolkovo Institute of Science and Technology (Nobelya ul. 3, Moscow, Skolkovo, 143026 Russian Federation).

Zhukov R.A. — post-graduate student, Skolkovo Institute of Science and Technology (Nobelya ul. 3, Moscow, Skolkovo, 143026 Russian Federation).

Teteryukov D.O. — Cand. Sc. (Eng.), Assist. Professor, Skolkovo Institute of Science and Technology (Nobelya ul. 3, Moscow, Skolkovo, 143026 Russian Federation).

Mozgovoy M.V. — Deputy Director for Methodological Work of the Head Training and Research and Methodological Center, Research and Methodological Center for the Disabled, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Volkov A.V. — Director of the Resource Training and Methodological Center for the Education of the Disabled, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Please cite this article in English as:

Suvorov D.A., Zhukov R.A., Teteryukov D.O., Mozgovoy M.V., Volkov A.V. Recognition of Russian Banknote Nominal Values by Mobile Devices for Blind People. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Priborostr.* [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.], 2018, no. 1, pp. 94–104 (in Russ.). DOI: 10.18698/0236-3933-2018-1-94-104