Применение глубоких свёрточных нейронных сетей для визуализации препятствий на взлётно-посадочной полосе

В.В. Князь^{1,2}, О.В. Выголов¹, В.В. Федоренко¹, В.С. Горбацевич¹, В.А. Мизгинов¹, Д.Г. Клевцов¹ vl.kniaz@gosniias.ru|o.vygolov@gosniias.ru|vfedorenko@gosniias.ru|gvs@gosniias.ru vl.mizginov@gosniias.ru|k_dm@gosniias.ru

¹125319, Москва, ул. Викторенко, 7;

²Московский физико-технический институт (государственный университет), Москва, Россия

Контроль чистоты поверхности взлётно-посадочной полосы является важным фактором обеспечения безопасности воздушного движения. Современный уровень развития датчиков и методов обработки поступающих от них данных обеспечивает необходимые предпосылки для решения задачи обнаружения и активной визуализации препятствий по данным тепловизионных камер в автоматическом режиме, что позволяет повысить эффективность уборки взлётно-посадочной полосы. Перспективным средством анализа мультиспектральных видеопоследовательностей являются глубокие свёрточные нейронные сети. В данной статье рассматривается модифицированная архитектура глубокой свёрточной сети SqueezeNet для задач автоматического выделения препятствий по видеопоследовательностям видимого и дальнего инфракрасного диапазонов.

Ключевые слова: глубокие свёрточные нейронные сети, тепловизионные изображения, обнаружение препятствий, взлётно-посадочная полоса, семантическая сегментация

Deep convolutional neural network for visualisation of foreign objects on a runway

V.V. Kniaz^{1,2}, O.V. Vygolov¹, V.V. Fedorenko¹, V.S. Gorbatsevich¹, V.A. Mizginov¹, D.G. Klevtsov¹ vl.kniaz@gosniias.ru|o.vygolov@gosniias.ru|vfedorenko@gosniias.ru|gvs@gosniias.ru vl.mizginov@gosniias.ru|k_dm@gosniias.ru

¹State Res. Institute of Aviation Systems (GosNIIAS), 125319, 7, Victorenko str., Moscow, Russia; ²Moscow Institute of Physics and Technology (MIPT), Russia

The presence of foreign objects on airport runways poses a significant threat to the safety of air travel. Infrared camera based runway monitoring systems for automatic detection and visualisation of foreign objects are highly demanded nowadays. Deep neural networks have recently became a powerful instrument for analysis of multispectral image sequences. This paper is focused on the development of an new deep neural network architecture for automatic detection of foreign objects on a runway. The architecture is based on the SqueezeNet network. The new network performs detection using a pair of images captured in visible and far infrared ranges.

Keywords: deep convolutional neural networks, thermal vision, obstacle detection, runway, semantic segmentation

1. Введение

Перспективным способом уборки взлётно-посадочной полосы (ВПП) являются колёсные роботы, оснащённые тепловизионными системами технического зрения. Управление роботом осуществляется удалённо из диспетчерского пункта. Оператор периодически производит осмотр поверхности полосы с использованием системы технического зрения мобильного робота.

Восприятие тепловизионных изображений требует навыков в силу значительных отличий от изображений видимого диапазона. Это обуславливает актуальность предварительной обработки тепловизионных изображений для акцентированной визуализации препятствий. Данная задача может быть решена с помощью методов семантической сегментации. Анализ последних работ в данной области [5–7] показывает, что наиболее перспективным подходом для решения задач семантической сегментации являются глубокие свёрточные нейронные сети. В данной работе рассматривается разработка новой архитектуры глубокой свёрточной нейронной сети для выделения препятствий на ВПП. На вход сети поступает многоспектральное изображение, включающее каналы в видимом и инфракрасном диапазонах. На выходе сеть выдаёт меточное изображение, на котором визуализируются два класса объектов: фон и препятствие. Разработанная архитектура основывается на архитектуре сети SqueezeNet [8]. В отличие от задачи классификации объекта на изображении, при решении задачи семантической сегментации требуется предсказать тип объекта в каждом пикселе исходного изображения. Для восстановления исходного разрешения изображения к архитектуре сети SqueezeNet были добавлены два слоя деконволюции.

По сравнению с альтернативными методами выделения препятствий на ВПП предлагаемый подход обладает рядом преимуществ. Во-первых, свёрточные нейронные сети позволяют производить обучение без насыщения: при поступлении новых изображений в обучающую выборку качество распознавания препятствий постоянно растёт. Во-вторых, в отличие от систем, основанных на выделении препятствий на основе модели движения камеры [11, 12], предлагаемый подход обеспечивает выделение препятствий при произвольной траектории движения.

2. Постановка задачи

Рассматривается задача в следующей постановке: колёсный робот, оснащённый многоспектральной системой технического зрения (СТЗ) FLIR ONE (рис. 1) производит обзор поверхности ВПП. Требуется осуществить обработку изображений, поступающих от СТЗ, для акцентированной визуализации типовых препятствий (выделение контуров, маркирование цветом и т.п.). Нахождение визуальных контуров препятствия производится с помощью глубокой свёрточной нейронной сети. Таким образом, постановка задачи сводится к семантической сегментации многоспектральных изображений на два класса: фон и препятствие.



Рис. 1. Мобильный робот с многоспектральной камерой FLIR ONE

Целевая функция. Пусть на вход сети поступает многоспектральное изображение $\mathbf{X} \in \mathbb{R}^{H \times W \times 3}$. Требуется обучить целевую функцию $\hat{\mathbf{Y}} = \mathcal{F}(\mathbf{X})$, которая производит отображение комплексированного изображения в меточное изображение $\mathbf{Y} \in \{0, 1\}^{H \times W}$, где H, W – размеры изображения, а метки 0 и 1 обозначают, соответственно, фон и препятствие. Для достижения равномерной сходимости обучения при обучении используется многомодальная функция потерь $L_c l$ [10], которая задаётся уравнением

$$L_{c}l(\hat{\mathbf{Y}},\mathbf{Y}) = -\sum_{h,w} v(\mathbf{Y}_{h,w}) \sum_{q} \mathbf{Y}_{h,w,q} \log(\hat{\mathbf{Y}}_{h,w,q})$$
(1)

3. Архитектура сети

Анализ последних работ в области семантической сегментации[3, 5–7] показывает, что наиболее высокой точности классификации позволяют достичь алгоритмы на основе глубоких свёрточных нейронных сетей. В отличие от традиционных архитектур сетей, используемых для классификации изображений, в архитектурах сетей для семантической сегментации, как правило, отсутствуют полносвязные слои. Также в данных сетях широко используются слои деконволюции.



Рис. 2. Архитектура сети

Для обучения целевой функции $\hat{\mathbf{Y}}$ требовалось разработать архитектуру глубокой свёрточной нейронной сети для семантической сегментации. Для достижения вычислительной эффективности за основу архитектуры была взята сеть SqueezeNet [8]. Основным преимуществом сети SqueezeNet является выигрыш в производительности в 50 раз по сравнению с сетью AlexNet[4], при сохранении сопоставимой точности классификации. Для обеспечения вычислительной эффективности произведено уменьшение размеров фильтров свёртки с 3×3 на 1×1 . Благодаря этому число обучаемых параметров сократилось в 9 раз. Таким образом, сеть строится из однотипных модулей, называемых «fire module». Для обеспечения устойчивого выделения препятствий в архитектуру разрабатываемой сети (рис. 2) был внесён ряд изменений. Во-первых, для восстановления пространственного разрешения входного изображения были добавлены два слоя деконволюции. Вовторых, для уменьшения количества параметров удалён слой субдискретизации (global avgpool).

4. Построение обучающей выборки

Обучающая выборка для семантической сегментации состоит из пар геометрически выровненных многоспектральных изображений X и меточных изображений Y. Комплексирование видимого и инфракрасного изображений осуществлялось путём замены красного канала исходного изображения видимого диапазона на изображение инфракрасного диапазона.

Для построения истинных меточных изображений для обучающей выборки использовался метод оценки траектории камеры, предложенный в [13]. Примеры исходных изображений, а также компенсированного и меточного изображений приведены на рисунке 3. Съёмка исходных изображений осуществлялась с помощью прототипа мобильного колёсного робота для уборки ВПП [1, 2].



(с) Многоспектральное изображение **X**

(d) Меточное изображение **Y**

Рис. 3. Примеры изображений из обучающей выборки. Условное препятствие (разводной ключ) расположено в центре кадра

Прототип мобильного робота для уборки ВПП. Для подготовки тестовой выборки использовался колёсный робот, оснащённый многоспектральной системой технического зрения. Робот построен на базе колёсной платформы «Hercules» и одноплатного компьютера «Raspberry Pi» (рисунок 1). В качестве многоспектральной системы технического зрения использовалась многоспектральная камера «FLIR ONE», подключённая к смартфону. Разрешение камеры в видимом диапазоне составляет 640х480 пиксел, в инфракрасном - 160х120 пиксел. Параметры камеры FLIR ONE приведены в таблице 1.

Характеристики FLIR ONE	
Диапазон температур	-20° C to 120° C
Bec	29 грамм
Размеры (Д х Ш х В)	72 х 26 х 18 мм
Чувствительность	0.1° C
ТВ камера	640x480 пикс.
ИК камера	160x120 пикс.
Поле зрения	46°x35°

Таблица 1. Характеристики камеры FLIR ONE

Управлением роботом осуществляется с персонального компьютера по каналу сети WiFi. Для управления роботом был разработан специальный протокол и приложения клиент/сервер для «Raspberry Pi» и персонального компьютера.

5. Обучение сети и анализ результатов

Обучение разработанной архитектуры сети производилось в среде NVIDIA DIGITS с использованием графического процессор Titan X PASCAL.

Тестирование обученной сети проводилось на тестовой выборке, объёмом в 300 пар изображений. Для определения качества работы сети использовалась бинарнарная мера сходства Intersection over Union (IU), определяемая соотношением:

$$IU = \frac{A_I}{A_U},\tag{2}$$

где A_I – площадь пересечения области препятствия на построенном меточном изображении и на эталонном меточном изображении, A_U – площадь объединения данных областей. Как правило, значение IU > 50.0% считается показателем приемлемого качества сегментации [9]. Примеры полученных меточных изображений представлены на рисунке 4. Значения IU для различных условий приведены в таблице 2.

Условия	Intersection over Union
Тень	62.6
Солнце и тень	50.1
Мокрый асфальт	49.3

Таблица 2. Значения IU для различных условий



(a) Комплексирование **(b)** I

(b) Результат работы сети

Рис. 4. Пример работы сети

6. Выводы

В работе предложен метод визуализации препятствий на поверхности ВПП с использованием глубокой свёрточной нейронной сети. Метод предназначен для решения задачи помощи оператору мобильного колёсного робота в обнаружении посторонних препятствий на поверхности ВПП. Он позволяет повысить эффективность обнаружения препятствий в ночное время и в условиях ограниченной видимости.

Для реализации метода разработана архитектура сети для выделения визуальных контуров препятствий. Произведено обучение и тестирование сети. Для обучения сети использовалась обучающая выборка объёмом 4000 пар комплексированных и меточных изображений. Съёмка изображений для обучающей и тестовой выборок производилась с помощью колёсного робота, оснащённого многоспектральной камерой FLIR ONE. Выборка включает снимки условных препятствий (разводной ключ, детали конструкции, предметы багажа и т.п.), полученные в различных условиях: прямой солнечный свет, граница света и тени, мокрый асфальт. Истинные меточные изображения были построены с использованием программы трёхмерного моделирования. Обучение сети производилось в среде NVIDIA DIGITS.

Тестирование разработанной сети производилось на независимой выборке объёмом в 300 пар изображений. Для определения точности выделения препятствий использовалась метрика Intersection over Union. Анализ полученных результатов показал, что сеть успешно производит выделение препятствия во всех рассмотренных условиях. Средняя точность выделения контуров препятствия по метрике Intersection over Union составила 54%.

7. Литература

- Kniaz V V. Fast instantaneous center of rotation estimation algorithm for a skied-steered robot // SPIE Optical Metrology. — 2015. — Jun. — Vol. 9528. — P. 95280L–95280L–11.
- [2] Kniaz V V. Real-time optical flow estimation on a GPU for a skied-steered mobile robot. — 2016. — Apr. — P. 989706–989706–12.
- [3] Kniaz V. V., Gorbatsevich V. S., Mizginov V. A. THERMALNET: A DEEP CONVOLUTIONAL NETWORK FOR SYNTHETIC THERMAL IMAGE GENERATION // ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2017. — Vol. XLII-2/W4. — P. 41–45. — URL: http://www.intarch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W4/41/2017/.
- [4] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems. – 2012.

- [5] Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015.
- [6] Long J., Shelhamer E., Darrell T. Fully Convolutional Models for Semantic Segmentation // CVPR 2015, and PAMI 2016. – 2016.
- [7] Semantic image segmentation for information presentation in enhanced vision / Oleg V. Vygolov, Vladimir S. Gorbatsevich, Nikita A. Kostromov et al. // Proc. SPIE. 2017. Vol. 10197. P. 101970H–101970H–8. URL: http://dx.doi.org/10.1117/12.2262507.
- [8] SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size / Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf et al. // CoRR. 2016 ____ Vol. abs/1602.07360. URL: ____ http://arxiv.org/abs/1602.07360.
- [9] The Pascal Visual Object Classes (VOC) Challenge / Mark Everingham, Luc Van Gool, Christopher K I Williams et al. // International Journal of Computer Vision. — 2009. — Vol. 88, no. 2. — P. 303–338.
- [10] Zhang Richard, Isola Phillip, Efros Alexei A. Colorful Image Colorization. // ECCV. — 2016. — Vol. 9907, no. Chapter 40. — P. 649–666.
- [11] Zheltov S, Sibiryakov A. Method of 3d-object detection based on orthophoto difference analysis // International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B3. Amsterdam 2000. – 2000. – P. 1041–1047.
- [12] Бусурин В. И., Князь В. В. Автоматическое обнаружение препятствий на взлётно-посадочной полосе средствами технического зрения. — Труды МАИ, 2014. — Т. 81.
- [13] Князь В. В., Мизгинов В. А., Горбацевич В. С. Создание модельных изображений инфракрасного диапазона и оценка их визуального качества с использованием свёрточных нейронных сетей. — Научная визуализация, 2016. — Т. 8. — С. 67–79.

Об авторах

Князь Владимир Владимирович, инженер ГосНИИАС, магистрант ФУПМ МФТИ. Его e-mail vl.kniaz@gosniias.ru.

Выголов Олег Вячеславович, к.т.н., начальник лаборатории ГосНИИАС. Его e-mail o.vygolov@gosniias.ru.

Федоренко Вадим Вадимович, инженер ГосНИИАС. Его e-mail vfedorenko@gosniias.ru.

Горбацевич Владимир Сергеевич, начальник лаборатории ГосНИИАС. Его e-mail gvs@gosniias.ru.

- Мизгинов Владимир Андреевич, инженер ГосНИИАС. Его e-mail vl.mizginov@gosniias.ru.
- Клевцов Дмитрий Геннадьевич, инженер ГосНИИАС. Его e-mail k_dm@gosniias.ru.