

УДК 004.457

DOI: 10.17223/19988605/45/9

М.С. Сорокин, Г.Б. Захарова**ИНТЕЛЛЕКТУАЛЬНЫЙ ПРОГРАММНЫЙ ИНСТРУМЕНТАРИЙ
ДЛЯ ПОДГОТОВКИ НОВОСТНЫХ ТЕЛЕВИЗИОННЫХ ПРОГРАММ**

На основе структуризации бизнес-процессов телерадиокомпании ГТРК «Урал» разработаны подход и средства автоматизации, которые позволили повысить эффективность производства новостных телевизионных программ почти на 70%. Решена проблема необходимости повторного ввода данных за счет конструктора интерфейсов, автоматизирован процесс наложения графики на видео, создан обработчик роликов для приведения звука и видео к единым стандартам. Для архивации видео и эффективного поиска в архиве разработана система, которая выполняет автоматическое распознавание видеоизображений с помощью нейронных сетей на основе машинного обучения.

Ключевые слова: TV-технологии; автоматизация бизнес-процессов; конструктор интерфейсов; наложение графики на видео; нейронные сети; распознавание видеоизображений.

Подготовка программ в современной телевизионной компании отличается большой сложностью, трудоемкостью, динамичностью и слабой степенью автоматизации. При разработке медийной продукции задействуется большое количество сотрудников – режиссеры, редакторы, корреспонденты, художники-графики, монтажеры. Все они взаимодействуют в сложной информационной среде, многие действия не формализованы, человеческий фактор влияет на время и качество производства. Только структуризация и автоматизация бизнес-процессов может повысить эффективность работы и улучшить общее качество выпускаемого контента.

Универсальных решений по автоматизации работ на телевидении, в частности для производства новостных программ, на сегодняшний день не существует. Более того, процессы являются достаточно закрытыми, узконаправленными, требуют больших трудозатрат. В качестве аналогов автоматизации можно назвать систему производства новостей NewsNet [1], в которой реализованы средства работы с лентами в едином пространстве для обмена информацией и получения доступа к медиа архиву телеканала. Открытая универсальная система Cinegy News [2] для автоматизированного производства и выпуска новостей содержит интегрированные между собой модули записи входящих новостных потоков и поступающих видеосюжетов, верстки сюжетов и выпусков новостей. Высокоавтоматизированным является также новостное производство в среде Media Backbone Hive от Sony [3]. Однако интеллектуальные функции в названных ресурсах не реализованы.

В данной работе на примере одной из ведущих телевизионных компаний Свердловской области – ГТРК «Урал» – показано, как многие стратегически важные бизнес-процессы по подготовке новостных программ были сокращены по времени в десятки раз с улучшением качества выполнения задач. Работа выполнялась на основе тщательного анализа всех выполняемых работ по производству медиа продукции по методике [4]. В результате была построена подробная схема бизнес-процессов аналогично тому, как это сделано в работе [5]. Все процессы были классифицированы на допускающие полную или частичную автоматизацию, а также выделены неавтоматизируемые функции. Схема не приводится в силу большого объема, были выделены, реализованы и описаны далее автоматизированные процессы.

Выполненная автоматизация бизнес-процессов в TV-компании построена на принципах модульности, интерактивности, масштабируемости, что существенно сократило число ошибок при подготовке новостных программ и значительно повысило эффективность рабочего процесса.

1. Конструктор интерфейсов для наложения текстов на видеоролики

Компания ГТРК «Урал» производит новостные телевизионные программы для двух телеканалов: «Россия 1» – 12 выпусков в день и «Россия 24» – 2 выпуска в день, базовая графика на каналах имеет различный формат. Для сюжетов приходится готовить множество информационных роликов, ролики для данных телеканалов также имеют ряд существенных отличий, для их создания используются разные проекты и, соответственно, разные шаблоны. Имеется порядка 10 видов информационных роликов, к примеру тексты с фото и без них, трансляция звука с фото и без фото, анонсы, промо и пр. Таким образом, объем и разнообразие производимой ежедневно медийной продукции огромны, на ее производство уходит в среднем до пяти человеко-часов в день, многие операции носят рутинный характер и делаются вручную.

Для автоматизации данного вида работ, в частности ввода текста в многообразные шаблоны роликов, был разработан конструктор интерфейсов, позволивший унифицировать действия пользователя. Инструмент позволяет для ввода данных создавать интерфейсы из типовых групп и панелей, содержащих все элементы графического пользовательского интерфейса. Это чекбокс для управления параметром с двумя состояниями, радиокнопка для выбора одной опции из predetermined набора, выпадающий список для выбора одного из нескольких заранее определенных значений параметра, надпись (static text), поле ввода (dynamic text) для ввода небольшого объема текста без переноса строк, вкладка (tabs), которая позволяет в окне приложения переключаться между несколькими формами, кнопка, «нажатие» на которую приводит к некоторому действию из заготовленных шаблонов или выполнению пользовательского скрипта.

При использовании данных интерфейсов пользователь получает возможность многократно применять единообразные удобные формы для ввода данных в проект (ранее ввод не был унифицирован), а также импортировать и экспортировать их через XML.

Технология реализации данного инструмента опирается на возможности программы Adobe After Effects, предназначенной для редактирования видео и динамических изображений, применения цифровых видеоэффектов и пр. Данная программа является де факто промышленным стандартом для производства медийной продукции на телевидении.

Конструктор реализован в среде Adobe ExtendScript Toolkit CC, которая обеспечивает доступ к многофункциональным средствам разработки и выполнения скриптов JavaScript для работы напрямую с приложениями Adobe, поддерживающими сценарии. Сценарий – это последовательность команд, предписывающих приложению выполнить последовательность операций. Сценарии можно использовать для автоматического выполнения повторяющихся задач, сложных вычислений, а также для выполнения некоторых функций, непосредственно не представленных в графическом пользовательском интерфейсе. Сценарий, который использует Adobe, называется ExtendScript (JSX), он является диалектом JavaScript или, точнее, EcmaScript.

Конструктор представляет собой JavaScript файл, который вызывается напрямую из After Effects. Для этого необходимо открыть проект, затем меню File, Scripts, Form_constructor.jsx. Конструктор работает в два этапа: сначала пользователь выбирает поля ввода и элементы управления, а далее производится связка формы и проекта After Effects через управляющий слой. Из конструктора формы сохраняются в файлы JSX для дальнейшего использования. Вызвать форму можно из программы After Effects через меню File. Формы поддерживают возможность выгрузки и загрузки данных в файл XML, что позволяет использовать одни данные без повторного ввода в нескольких проектах Adobe After Effects, значительно повышая производительность и понижая риск возникновения ошибки. Производительность работ составила порядка трех человеко-часов вместо пяти.

2. Модуль подключения текстовых и графических данных к видеоформату

Существует ряд задач, где автоматизация ввода данных в текстовые поля недостаточна. Необходимо обеспечить сложную связь между полями в форме, содержащей данные, с возможностью

регулировать продолжительность отображения наложенных на видео графики и текста, а также задавать ключевые кадры в ролике. С этой целью разработан инструмент, достаточно сложный в настройке и требующий в использовании понимания базовых принципов программирования и логики.

Одна из решаемых таким образом задач – подготовка роликов метеопрогноза. Каждый выпуск новостей содержит метеопрогноз, над созданием которого работают художники компьютерной графики. На создание качественного шаблонного проекта погоды уходит до трех месяцев, сюда входят съемка, отбор, обработка видеоматериала, разработка компьютерной графики и анимации. Прогноз погоды представляет собой текстовые и графические элементы, наложенные поверх фонового видеоролика. Каждый день создается четыре вида роликов: утренняя погода, погода сегодня, погода завтра и студийная погода.

Несмотря на кажущуюся простоту, эта повседневная работа требует большой концентрации и трудозатрат. Соответствующие бизнес-процессы с некоторыми отличиями в каждом из четырех видов прогноза содержали следующие действия: получение документа от гидрометцентра в формате Word, корректировка и приведение этого документа к единому виду, открытие проекта Adobe After Effects с шаблоном погоды, ввод данных из Word в Adobe After Effects, отправка на рендер, рендеринг и перекодирование полученного видеоролика с использованием кодека QuickTimeb в файловой транскодинг-системе Rhozet ProMedia Carbon, далее отправка в сетевое хранилище. На подготовку прогноза ежедневно уходило порядка четырех часов.

Для повышения эффективности работы был реализован модуль подготовки видеороликов метеопрогноза GeneratorXML на основе единого файла данных о погоде. Файл в формате XML, удобный для машинного чтения, содержит все необходимые данные для подготовки информационных заставок погоды в течение дня. Через эргономичный интерфейс данные вводятся однократно, без дублирования, с возможностью редактирования имеющихся, в том числе архивных, данных. Цепочка бизнес-процессов свелась при этом до двух действий: получение документа Word от гидрометцентра и ввод этих данных в разработанный модуль подготовки погоды. Затрачиваемое время – 20 минут.

Далее для изготовления ролика необходимо создать множество ключевых кадров, которые отражают любое изменение состояния объектов на видеоизображении. Для подключения анимации без необходимости создания множества ключевых кадров вручную был использован такой инструмент в Adobe After Effects, как выражения. Язык выражений основан на стандартном языке JavaScript. Выражение используется для оценки одиночного свойства слоя в конкретный момент времени и сообщает приложению о характеристике свойства. С помощью выражений можно создавать связи между свойствами слоя и использовать ключевые кадры конкретного свойства для создания динамической анимации других слоев.

Для автоматизации данного вида графики был добавлен шаблон формы на три дня, на неделю, формирование суточного прогноза по городу, реализована возможность связки полей друг с другом, возможность загружать данные из Интернета.

В результате автоматизации трудозатраты по подготовке видео с наложением данных от гидрометцентра сократились почти в четыре раза и составили немногим более часа.

3. Автоматизация обработки подготовленных видеороликов под оборудование

По стандартам телекомпании видеофайлы должны быть строго определенного формата с конкретными параметрами и соответствовать спецификации оборудования. Поскольку видео поступает в различных форматах, был реализован обработчик готовых роликов для перекодировки видеофайлов, приведения звука и видео к общепринятым стандартам.

Обработчик видео и звука представляет из себя конвертор на базе языка Python и библиотеки FFmpeg для работы с медиафайлами. Конвертор принимает на вход файлы в любом поддерживаемом формате и с любыми параметрами через заранее заданную папку. Эта папка становится специальной смотрящей (Watch Folder) папкой; любой файл, который оказывается в ней, попадает в обработку.

При попадании на конвертор видеофайл сначала сжимается до заданного разрешения, а также приводит форму пикселя к стандарту, в случае с компанией ГТРК «Урал» – PAL DV (720 × 576, Aspect ratio – 1,07). Далее идет работа с чересстрочностью, устанавливается нижний полукадр. Затем обрабатывается звук: устанавливается частота дискретизации Sampling rate 48,0 kHz, общий уровень производимого шума приводится к 23 LUFS (Loudness Unit relative to Full Scale – единица громкости относительно полной шкалы) в соответствии со стандартом R-128 Европейского вещательного союза (EBU).

На выходе конвертор возвращает файлы, соответствующие требованиям, с пометкой о соответствии. Настройка параметров обработки производится через простой графический интерфейс.

Время обработки файлов с применением конвертора сократилось почти вдвое: 17 минут вместо получаса.

4. Автоматическое формирование архива и распознавание изображений в видеофайлах

ТВ-компании в процессе деятельности создают огромное количество видео контента от видеосъемок до компьютерной графики, во многих случаях эти проекты подразумевают повторное использование. Организуются архивы с соответствующим штатом сотрудников, которые каталогизируют материалы и по запросам предоставляют доступ. В целях автоматизации данного вида деятельности разработана система, позволяющая структурировать и архивировать видеоматериал и выполнять его эффективный поиск на основе распознавания изображений в видеофайлах. Процесс распознавания выполняется посредством сверточной нейронной сети, в данном случае применялась открытая библиотека Keras [6], настроенная на работу с сетями глубокого обучения. В отличие от программы, описанной в работе [7], такой подход позволил распознавать более сложные сцены, но в рамках заданных классов. Реализация выполнена на языке Python [8].

В данной системе автоматизации поиска видео можно выделить три части:

- 1) поддержка видеоархива: фоновое приложение, которое при внесении в первичный архив нового видео выполняет его обработку по выделению ключевых кадров и оптимизацию;
- 2) распознавание изображений на видео с помощью нейронной сети;
- 3) собственно поиск, определяющий по запросу пользователя искомый объект на видео и возвращающий пути к найденным релевантным видео в архиве.

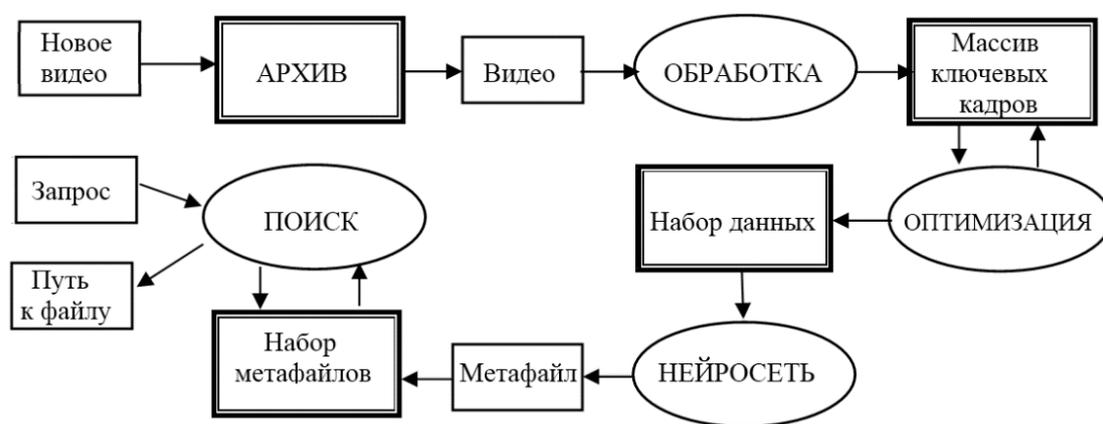


Рис. 1. Структура системы поддержки видеоархива и поиска видео

4.1. Поддержка видеоархива

Фоновое приложение получает на вход папку отслеживания. На фазе инициализации программа проверяет соответствие каждого видеофайла в папке с уже имеющимися метаданными к этим файлам.

При выявлении видеофайлов без метаданных, а также при появлении новых файлов, в папке создается поток для построения соответствующих данных. В потоке формируется дочерний процесс, который позволяет получить из видеофайла массив ключевых кадров. Набор ключевых кадров хранится во временной папке. Есть несколько способов отделения видео от шума, выделения релевантной информации из кадров: 1) взять все кадры – кадр за кадром; 2) брать каждый n -й кадр; 3) отсеивать пустые, повторяющиеся кадры, оставляя наиболее информативные. В данной системе используется третий подход. Массив изображений сокращается за счет удаления схожих между собой и прочих не содержащих полезной информации кадров.

4.2. Распознавание изображений на видео

На основе массива изображений готовится набор данных (batch) для анализа нейронной сетью. Изображения классифицируются тремя моделями: 1) выделение объекта на изображении; 2) определение места внутри или вне помещения; 3) распознавание времени года. Модель (model), основная структура данных Keras, – это способ организации слоев. Простейшим типом модели является последовательная модель (sequential model), линейная совокупность слоев, которая и используется в данном случае. Когда сеть содержит несколько скрытых слоев, она приобретает способность самообучаться и выделять ключевые признаки, которые наилучшим образом описывают входные данные. Отличительное свойство Keras, которое выделяет его из других фреймворков (например, TensorFlow), – это автоматический расчет размеров слоев: достаточно указать только размерность входного слоя, а Keras автоматически проинициализирует все остальные слои.

Для обучения каждой модели нейронной сети использовалось по три набора изображений: набор данных для обучения (порядка 20 000 изображений); набор данных для проверки (оценка качества обучения сети в процессе обучения – 10 000 изображений); набор данных для тестирования (оценка качества после завершения обучения – 5 000 изображений). Для загрузки изображений в Keras используются генераторы изображений (ImageDataGenerator), которые позволяют автоматически загружать изображения с диска в память компьютера, преобразовывать их в вид, необходимый Keras, и передавать в модель.

Первая модель сети отвечает за определение класса, к которому относится изображение по виду объектов на нем: человек, автомобиль, компьютер и т.д. (Label Detection). Keras позволяет не только обучать глубокие нейронные сети, но и содержит несколько готовых, предварительно обученных сетей, которые можно использовать, не тратя время и ресурсы на обучение, так называемый перенос обучения (transfer learning). Большая часть сетей предназначена для распознавания объектов на изображениях из краудсорсинговой базы данных ImageNet [9], содержащей более 14 млн изображений, привязанных к семантическому тезаурусу английских существительных WordNet. Для распознавания применялась VGG16 – сверточная сеть из 16 слоев. Сеть состоит из двух частей: первая часть выделяет характерные признаки в изображении, а вторая отвечает за классификацию объекта по этим признакам. На выходе сеть выдает вероятности того (в формате one hot encoding), что на изображении тот или иной класс объекта.

Вторая модель сети представляет собой бинарный классификатор, который определяет место съемки внутри или снаружи помещения (Indoor / Outdoor) и содержит 10 слоев (8 для выделения признаков и 2 для классификации). Принимает на вход изображение размером 224×224 пиксела с тремя каналами RGB. Использует оптимизатор Adam (adaptive moment estimation), сочетает в себе идею накопления движения и идею более слабого обновления весов для типичных признаков.

Третья модель сети различает время года съемки (Season). Модель состоит из 13 слоев выделения признаков сети VGG16 плюс 3 слоя классификации. Характеристики входного изображения как в предыдущем случае.

Данные по кадрам усредняются. Выявляются доминирующие и наиболее вероятные образы. Из полученных сведений формируется метафайл, который содержит: хэш файла, общую продолжительность видеофайла, количество релевантных кадров, снаружи / внутри помещения, время года, 10–200 наиболее вероятных объектов.

4.3. Поиск

Приложение поиска предоставляет интерфейс с возможностью фильтрации по ключевым характеристикам, полученным в результате работы второй и третьей моделей (зима / лето, улица / помещение), и поиск по меткам в метафайле. Программа объединяет сведения о файлах в заданной папке, после чего производит поиск по этим данным. Далее выводит превью соответствующих запросу видеофайлов в порядке релевантности с описанием других имеющихся образов в этих видеофайлах.

Данная система позволила существенно сократить время поиска в видеоархиве и оперативно получать доступ к видеозаписям независимо от количества файлов в архиве.

Заключение

Представленная интеллектуальная система автоматизации подготовки новостных телевизионных программ внесла множественные изменения в бизнес-процессы TV-компании ГТРК «Урал». Наряду с описанными четырьмя инструментами был разработан ряд других программ: для подготовки промо-роликов, изготовления инфографики и др. Программный код составил порядка 20 тыс. строк, были разработаны словари баз данных суммарным объемом более 16 тыс. строк.

После внедрения системы были проведены измерения времени, затраченного на производство медиапродукции до и после автоматизации по всем видам производства. Общая экономия времени в день составила 320 минут: вместо 557 минут стало уходить 236, т.е. 68% экономии. Улучшилось качество материалов, уменьшилось количество ошибок за счет удобных интерфейсов ввода данных и отсутствия дублирующих процессов.

ЛИТЕРАТУРА

1. Система производства новостей News Net. 2016. 19 июля // Advview : сайт о телевидении, автоматизации, оборудовании, маркетинге и рекламе. URL: https://advview.ru/cat_automation/nrcs/sistema-proizvodstva-novostej-newsnet/ (дата обращения: 14.08.2018).
2. Cinegy News TV Pack: как автоматизировать выпуск новостей на телеканале. 2017. 22 авг. // Advview : сайт о телевидении, автоматизации, оборудовании, маркетинге и рекламе. URL: https://advview.ru/cat_automation/nrcs/cinegy-news-i-tv-pack-kak-avtomatizirovat-vypusk-novostej-na-telekanale/ (дата обращения: 14.08.2018).
3. Новостное производство с Media Backbone Hive от Sony. 2016. 10 окт. // Advview : сайт о телевидении, автоматизации, оборудовании, маркетинге и рекламе. URL: https://advview.ru/cat_automation/nrcs/novostnoe-proizvodstvo-s-media-backbone-hive-ot-sony/ (дата обращения: 14.08.2018).
4. Волков О.Н. Стандарты и методологии моделирования бизнес-процессов // Связьинвест : корпоративное издание. 2005. № 7. С. 20–24.
5. Быков В.В. Информационные технологии и автоматизация ТВ-производства // Т-Comm: телекоммуникации и транспорт. 2012. Т. 6, № 9. С. 46–49.
6. Keras code and weights files for popular deep learning models. URL: <https://github.com/fchollet/deep-learning-models> (accessed: 12.08.2018).
7. Голубинский А.Н., Толстых А.А. Распознавание объектов на телевизионных изображениях с использованием аппарата сверточных нейронных сетей // Вестник Воронежского института МВД России. 2017. № 1. С. 71–81.
8. Chollet F. Deep Learning with Python. Manning Publications? 2017. 384 p.
9. ImageNet. URL: <http://image-net.org/about-overview> (accessed: 12.08.2018).

Поступила в редакцию 22 июня 2018 г.

Sorokin M.S., Zakharova G.B. (2018) THE INTELLIGENT SOFTWARE TOOLKIT FOR PRODUCTION OF NEWS TV-PROGRAMS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 45. pp.76–82

DOI: 10.17223/19988605/45/9

Based on the structuring of business processes of the TV and Radio Company "GTRK Ural", automation tools have been developed that have significantly increased the efficiency of the production of news programs.

The designer of interfaces was developed and introduced, due to what the problem of necessity of repeated data input is solved. The constructor is a JavaScript file that is called directly from Adobe After Effects and works in two stages: the user constructs input

fields and interface controls, then the Adobe After Effects form and project are linked through the control layer, saving to JSX files and further use.

The management tools of the already prepared media products have been implemented. So, for the development and filling of computer graphics templates in the preparation of weather forecast videos, automation technology was used based on the expressions and scenarios of the Adobe After Effects program and the toolkit for overlaying graphics on the video image was created. The weather forecast generator module generates a common weather data file in XML format based on hydrometeorological information. The file, convenient for machine reading, contains all the necessary data for the preparation of weather information posters during the day. Through the ergonomic interface, data are entered once, without duplication, with the ability to edit existing, including archived, data. At the same time, the chain of business processes instead of six actions before automation was reduced to two actions: obtaining a Word document from the hydrometeorological center and entering this data into the developed module, which takes about 20 minutes a day instead of the previous 4 hours.

The converter for ready-made clips has been created to bring sound and video to unified standards, working through a special interface written in Python.

With a view to efficiently structuring and archiving video material and subsequent search in the archive, it was developed the system that allows to automate processes of recognizing the contents of video files using neural networks based on machine learning. The recognition process is performed by means of a convolutional neural network, an open Keras library was used, tuned to work with deep learning networks. Images are classified according to three models: Label Detection, a VGG16 convolutional network of 16 layers was used for recognition; the definition of a place inside or outside the room (Indoor / Outdoor); Season recognition, the model consists of 13 layers of VGG16 network feature highlighting plus 3 layers for classification.

After the introduction of the system in the working process, the time spent on the production of media products before and after automation was measured for all types of production. The total time savings per day was 320 minutes: instead of 557 minutes 236 were spent, i.e., 68% time savings. The quality of materials was improved, the number of errors was reduced due to convenient data entry interfaces and the absence of duplicate processes.

Keywords: TV-technology; overlay graphics on video; neural networks; recognition of video images.

SOROKIN Mikhail Sergeevich (Ural Federal University, Ekaterinburg, Russian Federation).

E-mail: sjfh@yandex.ru

ZAKHAROVA Galina Borisovna (Candidate of Technical Sciences, Associate Professor, Ural State University of Architecture and Art, Ural Federal University, Ekaterinburg, Russian Federation).

E-mail: zgb555@gmail.com

REFERENCES

1. Advview.ru. (2016a) *News Net news production system*. 19th July. [Online] Available from: https://advview.ru/cat_automation/nrcs/sistema-proizvodstva-novostej-newsnet/. (Accessed: 14th August 2018). (In Russian).
2. Advview.ru. (2017) *Cinegy News TV Pack: kak avtomatizirovat' vypusk novostey na telekanale* [Cinegy News TV Pack: how to automate the news release on the channel]. 22nd August. [Online] Available from: https://advview.ru/cat_automation/nrcs/cinegy-news-i-tv-pack-kak-avtomatizirovat-vypusk-novostej-na-telekanale/. (Accessed: 14th August 2018).
3. Advview.ru. (2016) *Novostnoye proizvodstvo s Media Backbone Hive ot Sony* [News production with Media Backbone Hive from Sony]. 10th October. [Online] Available from: https://advview.ru/cat_automation/nrcs/novostnoe-proizvodstvo-s-media-backbone-hive-ot-sony/. (Accessed: 14th August 2018).
4. Volkov, O.N. (2005) Standarty i metodologii modelirovaniya biznes-protsessov [Standards and methodologies for modeling business processes]. *Svyaz'invest*. 7. pp. 20–24.
5. Bykov, V.V. (2012) Information technologies and automation of TV production. *T-Comm: telekommunikatsii i transport – T-Comm: Telecommunications and Transport*. 6(9). pp. 46–49. (In Russian).
6. GitHub. (n.d.) *Keras code and weights files for popular deep learning models*. [Online] Available from: <https://github.com/fchollet/deep-learning-models>. (Accessed: 12th August 2018).
7. Golubinsky, A.N. & Tolstykh, A.A. (2017) Recognition of objects on television images using the apparatus of convolutional neural networks. *Vestnik Voronezhskogo instituta MVD Rossii – Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia*. 1. pp. 71–81. (In Russian).
8. Chollet, F. (2017) *Deep Learning with Python*. Manning Publications.
9. *ImageNet*. (n.d.) [Online] Available from: <http://image-net.org/about-overview> (Accessed: 14th August 2018). (In Russian).