# Реализации нейронной сети «ResNet» в задаче распознавания типов летательных аппаратов.

## Сокращения

|  |  |
| --- | --- |
| Сокращение | **Наименование** |
| НС | * Нейронная сеть |
| СНС | * Сверточная нейронная сеть |
| ЛА | * Летательный аппарат |
| РГО | * Распознавание графических образов |

## Введение

Создание многофункциональных локационных систем путем введения в их состав алгоритмических средств распознавания графических образов (РГО) различных типов летательных аппаратов (ЛА) является актуальной задачей, успешное решение которой непосредственно влияет на обороноспособность страны. В системах военного назначения распознавание типов ЛА необходимо для построения плана тактических и стратегических мероприятий и выработки алгоритма противодействия воздушным целям, исходя из анализа складывающейся воздушной обстановки. В настоящее время в связи с интенсивной теоретической разработкой и практическим применением технологий, маскирующих локационные характеристики целей, роль задачи распознавания типов ЛА существенно возрастает. Решение задачи классификации типа ЛА так же может применятся в системах «свой-чужой», и для предотвращения нанесения ударов по гражданским ЛА

Интенсивное развитие вычислительной техники в последнее время и использование новейших технологий в области РГО может существенно повысить точность решения задачи классификации типов ЛА, уменьшить требуемое время наблюдения ЛА до решения задачи классификации с заданной точностью.

Обзор современных подходов к задаче РГО дает однозначный ответ о лучших методах решения задачи РГО. Во всех международных конкурсах по классификации изображений нейронные сети глубокого обучения дают наилучший результат[14]. Из всего семейства НС следует выделить семейство сверточных НС (СНС) [15].

В то же время практически отсутствую упоминания о применении подобных сетей в задачах РГО в авиационной области. Данная работа призвана несколько устранить этот пробел.

## Цель

В данной работе автор ставит перед собой 2 основные цели. Первая цель: применить новейшие разработки в области СНС в задаче РГО, оценить результат и перспективность использования сложных СНС в подобных задачах. Препятствием к достижению этой цели является отсутствие (или недостаточность) обучающей выборки изображений ЛА разных типов. Действительно, лучшая СНС «ResNet», состоит из 50 слоев (та версия ResNet, которую я использовал), для обучения такой сети необходима выборка из десятков тысяч (как будет показано ниже) изображений каждого типа ЛА. Составить такую размеченную коллекцию изображений (размеченная в том смысле что каждому изображению необходимо указать класс объекта который на нем изображен), само по себе является отдельной ресурсоемкой задачей. Данная задача не может быть решена автоматически. Таким образом мы приходим к идеи генерировать изображения для обучения СНС, вместо того что бы брать реальные. Теперь цели номер два может быть сформулирована следующим образом: Получить ответ на вопрос, можно ли использовать сгенерированные изображения для обучения СНС, в частности СНС «ResNet».

## План работы

1. Генерация изображений ЛА нескольких разных типов.
2. Обучение (а точнее дообучение) СНС «ResNet» задаче классификации изображений этих типов ЛА.
3. Оценка результата классификатора по контрольным РЕАЛЬНЫМ (не сгенерированным) изображениям.

## Описание стенда

Данная работа выполнена в операционной среде Ubuntu14.04LTS, использован язык программирования python2.7. Библиотеки для работы с НС Theano, keras [7, 8]. Полностью код проекта доступен по ссылке [5]. Генерация изображений осуществлялась в редакторе «3ds Max 2016».

## Генерация изображений для обучения НС

Для генерации исходных изображений для НС были использован редактор «3ds Max 2016». В качестве исходных моделей были использованы общедоступные модели.

Для обучения были использованы модели F15, F16, F117, Миг29, Boeing747 [2](мига пока нету) каждая модель ЛА была представлена в 3 различных вариантах раскраски. Для генерации изображений были использованы 3 разных фоновых изображения. Каждый ЛА снимался с различных углов с одинакового расстояния. При генерации изображения использовались фильтры для наложения шумов на изображения (параметры шумов – размытие, зернистость). Из всего множества возможных изображений случайным образом для каждого типа ЛА было выбрано 6000 изображений случайным образом. Эти изображения, дополненные меткой типа ЛА использовались в качестве обучающего множества. Типичная серия сгенерированных изображений показана на рисунке Рисунок 1. Примеры серий сгенерированных изображений, использованные модели F16, F117Все изображения сгенерированные и использованные в данной работе доступны[6].







Рисунок 1. Примеры серий сгенерированных изображений, использованные модели F16, F117, F15.

## Обзор СНС «ResNet»

СНС «ResNet» [1] разработана группой специалистов из лаборатории Microsoft Research в 2015 году и показала лучший результат в международном конкурсе классификации изображений[14]. В данной статье приведена общая послойная архитектура «ResNet».

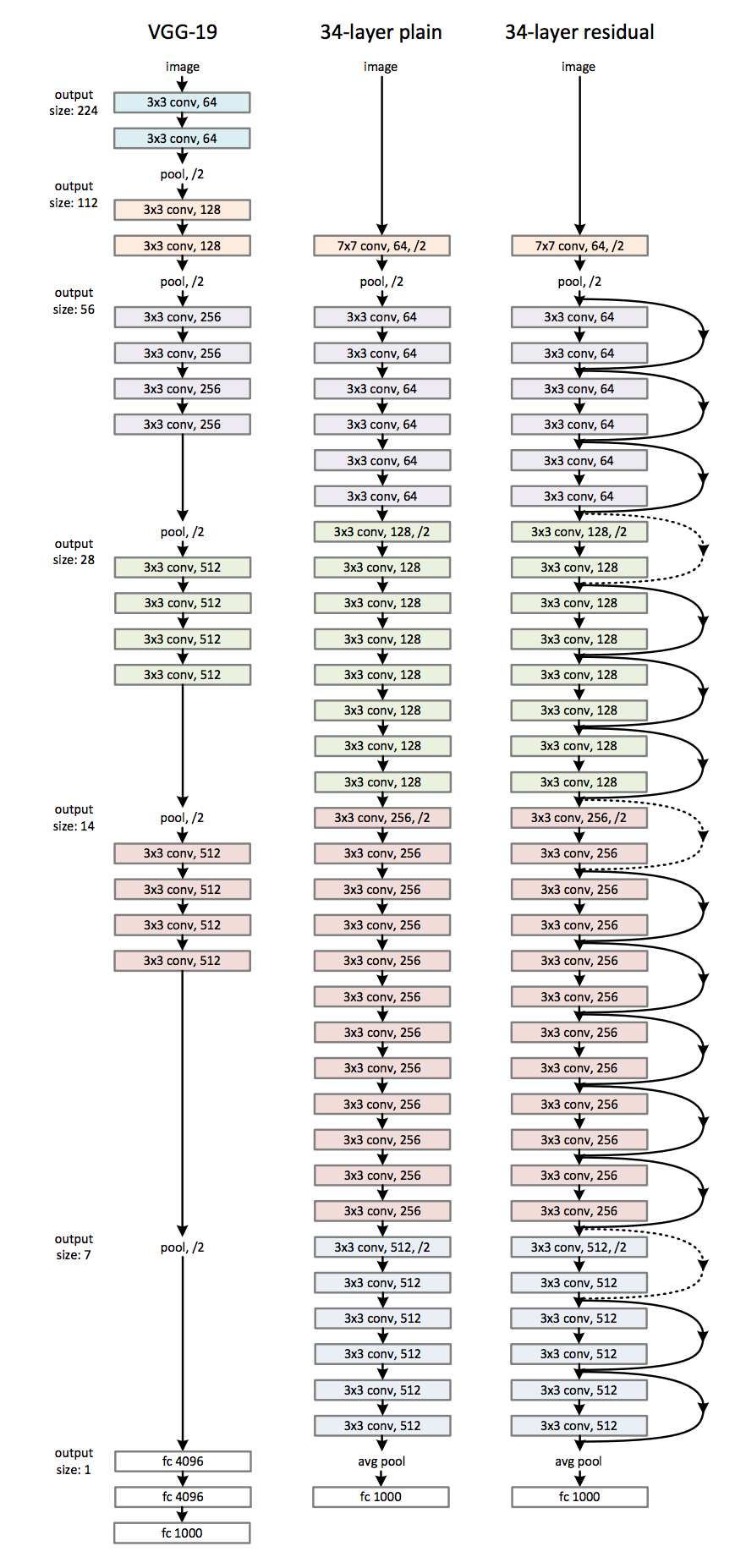


Рисунок . Архитектура НС «ResNet 34» Испавить на 50

Краткий обзор «ResNet»

Обзор первых слоев…

Пояснение что такое СНС и чем она выгодно отличается от полносвязных сетей…

Что добавляют связи «residual».

Количество нейронов в НС «ResNet» делают задачу обучения практически неподъёмной для домашних компьютеров. К счастью, существуют реализации этой НС уже обученные, т.е. вычислен вес каждого нейрона для одной из задач. В этой работе автор использовал первоначальные веса нейронов полученные в результате обучения НС на задаче классификации изображений соревнования «ImageNet»[15]. Т.о. не ставиться задача проведения полного обучения НС, но стоит задача проведения дообучения. Создатели «ResNet» не предполагали использование данной сети в задаче распознавания различных типов ЛА. В случае если обученной «ResNet» предложить классифицировать одно из изображений ЛА, представленных на рисунке Рисунок 1 то можно получить типовой ответ, представленный в таблицеТаблица 1.

|  |  |
| --- | --- |
| Класс объекта | Вероятность присутствия объекта на изображении |
| «warplane» | 0.8 |
| «aircraft» | 0.05 |
| … | <0.01 |

Таблица 1. Типовой ответ предобученной НС «ResNet» на запрос о классе ЛА (модель F15).

Т.е. «ResNet» имеет знание о классе «warplane» (боевой самолет), но не знает, на какие подклассы этот класс делиться. Прямое использование предобученной сети не представляется возможным.

## Исследование свойств предобученной НС для решения задачи классификации ЛА

Нам известно, что НС «ResNet» показывает высокие результаты на самых престижных соревнованиях по анализу изображений, будет логичным предположить, что даже без дообучения данная сеть может дать нам хорошее качество классификации. Для проверки этой гипотезы были выбраны 300 изображений F15 и F16. Т.к. НС не известны классы этих объектов, следовательно мы не можем использовать выход НС для оценки качества классификации, но можем использовать предпоследний слой, который имеет размерность 1000. Т.к. преобразование однослойной сети линейно, то, ясно, что НС до предпоследнего слоя выполнила работу по преобразованию входного изображения в вектор признаков размерностью 1000 таким образом, что бы все классы были линейно разделимы последним преобразованием (из предпоследнего слоя в последний). Т.о. мы можем использовать НС как инструмент преобразования изображение в вектор(1000). Получив для каждого из изображений вектор, размерностью 1000, мы можем воспользоваться методом главных компонент [11, 12] для отображения пространства 1000-мерной размерности в 2-х мерное пространство. Результат для случайных 600 ЛА (300 F15 и 300 F16) приведен на рисунке Рисунок 3.

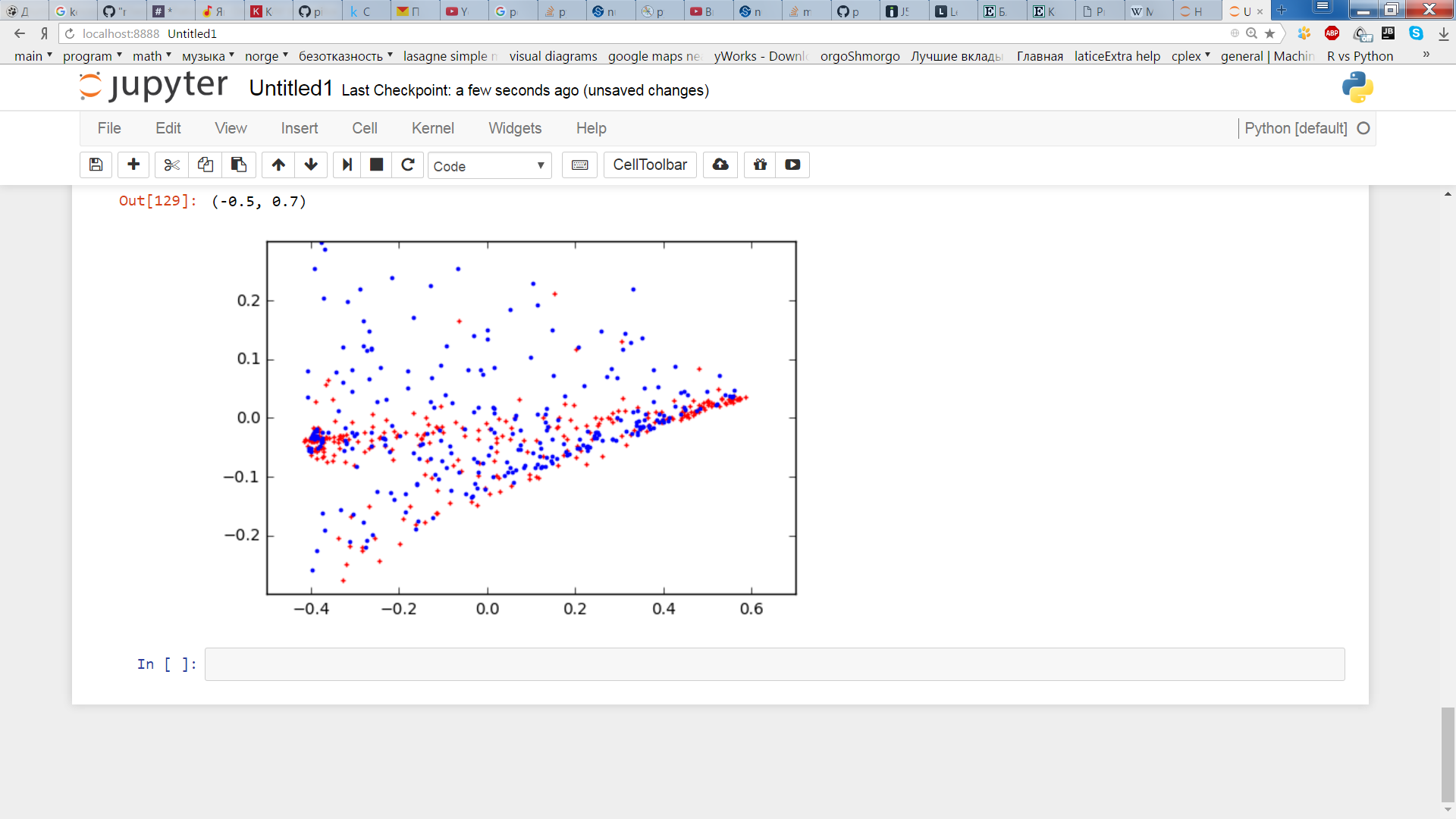


Рисунок 3. Проекция результатов выходного 1000-мерного слоя в 2-мерное пространство методом главных компонент. Красными ромбами обозначены выход для модели F16, синими точками – выход модели F15

После получения этого графика остается признать, что отличить модели F16 от F15 без некой доработки (адаптации) НС к настоящей задаче не представляется возможным.

## Адаптация НС для решения задачи классификации типов ЛА

Адаптация НС «ResNet» состоит из двух задач: адаптация архитектуры НС и дообучение НС (точная настройка весов нейронов для максимизации целевой функции решаемой задачи). В качестве последнего слоя я взял слой размерностью 5 (по числу исследованных в данной работе типов ЛА). Выход каждого слоя содержит вероятность принадлежности поданного на вход изображения соответствующему классу. Дообучение весов нейронов происходила путем показа НС всех сгенерированных изображений ЛА. Всего было проведено 80 эпох дообучения итоговые значения на вариационной выборке:

….

…потребные мощности. мощности Amazon, время дообучения.

## Численный результат

Реальные изображения на которых была проведена проверка результата НС приведены в хранилище [3]. Все изображения поочередно показывались «ResNet», полученный усредненный результат представлен в таблице ниже.

Пока у меня есть результат только по 2-м ЛА F15 и F16, НС была обучена по 9600 изображений (по 4800 каждый) проверялся на 15 изображениях из [3].

|  |  |  |
| --- | --- | --- |
| Тип ЛА | F15 | F16 |
|  | F15 | F16 |
| 6 изображений F15 | 80% | 20% |
| 6 изображений F16 | 22% | 78% |

этот результат уже не плох. На самом деле, я уверен при правильном подходе к дообучению НС этот результат можно довести до 99%!!! При получении этого результат (здесь решается бинарная задача классификации) сеть ни разу не ошиблась в прогнозе т.е. сеть могла дать результат с большей или меньшей увереностью, но НС ни разу не сказала, что F15 это более чем на 50% F16.

Типовые изображения:



В итоге должна получиться такая таблица:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Тип ЛА | F15 | F16 | F117 | Миг29 | Boeing 747 |
| F15 |  |  |  |  |  |
| F16 |  |  |  |  |  |
| F117 |  |  |  |  |  |
| Миг29 |  |  |  |  |  |
| Boeing 747 |  |  |  |  |  |

Таблица 2. Усредненные данные результатов классификации изображений НС «ResNet».

Зависимость качества классификации от объема обучающей выборки (количества сгенерированных моделей)

…

Зависимость качества классификации от количество классов ЛА

….

## Выводы и некоторые соображений которые возникли у меня во время выполнения данной работы

При увеличение количество классов качество классификации для каждого типа ЛА не проседает.

Первое представленная выше техника легко обобщается на задачу распознавания изображения на видео – из видео ряда необходимо сделать серию изображения, решить задачу классификации для каждого и далее, зная, что на видеокадрах представлен один и тот же объект воспользоваться теоремой о вероятности. Большое количество кадров даст нам высокую вероятность решения задачи классификации.

В рамках данной работы автор провел эксперимент с видео, доступным на youtube [10]. На видео снят полет самолета F16. На фоне городской застройки.

Результаты анализа видео.

… остальные соображения…

## Ссылки

1. ResNet main article. <https://arxiv.org/pdf/1512.00567v3.pdf>.
2. Модели F15, F16, Миг29, Boeng747 https://github.com/piero10/AircraftsRecognition/tree/master/models
3. Реальные - проверочные изображений ЛА.
4. <https://github.com/piero10/AircraftsRecognition/tree/master/images/realImages>
5. Хранилище исходного кода проекта. <https://github.com/piero10/AircraftsRecognition/tree/master>
6. Хранилище сгенерированных изображений ЛА. <https://github.com/piero10/AircraftsRecognition/tree/master/images/modelImages>
7. Библиотека низкого уровня для работы с НС - Theano. <http://deeplearning.net/software/theano/install_windows.html>
8. Библиотека высокого уровня для работы с НС - Keras. <https://keras.io/>
9. Imagenet. <http://image-net.org/challenges/LSVRC/2016/index>
10. Видео F16. <https://www.youtube.com/watch?v=aCmB3hz9UI0>
11. <http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>
12. <http://setosa.io/ev/principal-component-analysis/>
13. НС Resnet50 <https://github.com/KaimingHe/deep-residual-networks>
14. ResNet first place in classification task <http://image-net.org/challenges/LSVRC/2015/>
15. Сверточная НС <http://cs231n.github.io/convolutional-networks/>