# Адаптация нейросети глубокого обучения к 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[13]. Из всего семейства НС следует выделить семейство сверточных НС (СНС) [14]. В то же время практически отсутствую упоминания о применении нейронных сетей в задачах РГО в авиационной области. В тех работах, в которых упоминается о применении нейросетевого подхода к задаче РГО. Данная работа призвана несколько устранить этот пробел.

## Цель

В данной работе ставятся 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, СНС «ResNet», которая была выбрана в качестве основной, при написании данной статьи состоит из 34 слоев, для обучения такой сети необходима выборка из, по крайней мере тысяч, а лучше - десятков тысяч 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 [6, 7]. Полностью код проекта доступен по ссылке [4]. 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. Каждая модель ЛА была представлена в 3 различных вариантах раскраски. Раскраски одинаковые для разных типов ЛА. 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 для получения «реалистичного» изображения ЛА. 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, . Все изображения сгенерированные и использованные в данной работе доступны[5].





Рисунок 1. Примеры серий сгенерированных изображений, использованные модели F16, 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[13]. В данной статье приведена общая послойная архитектура «ResNet».

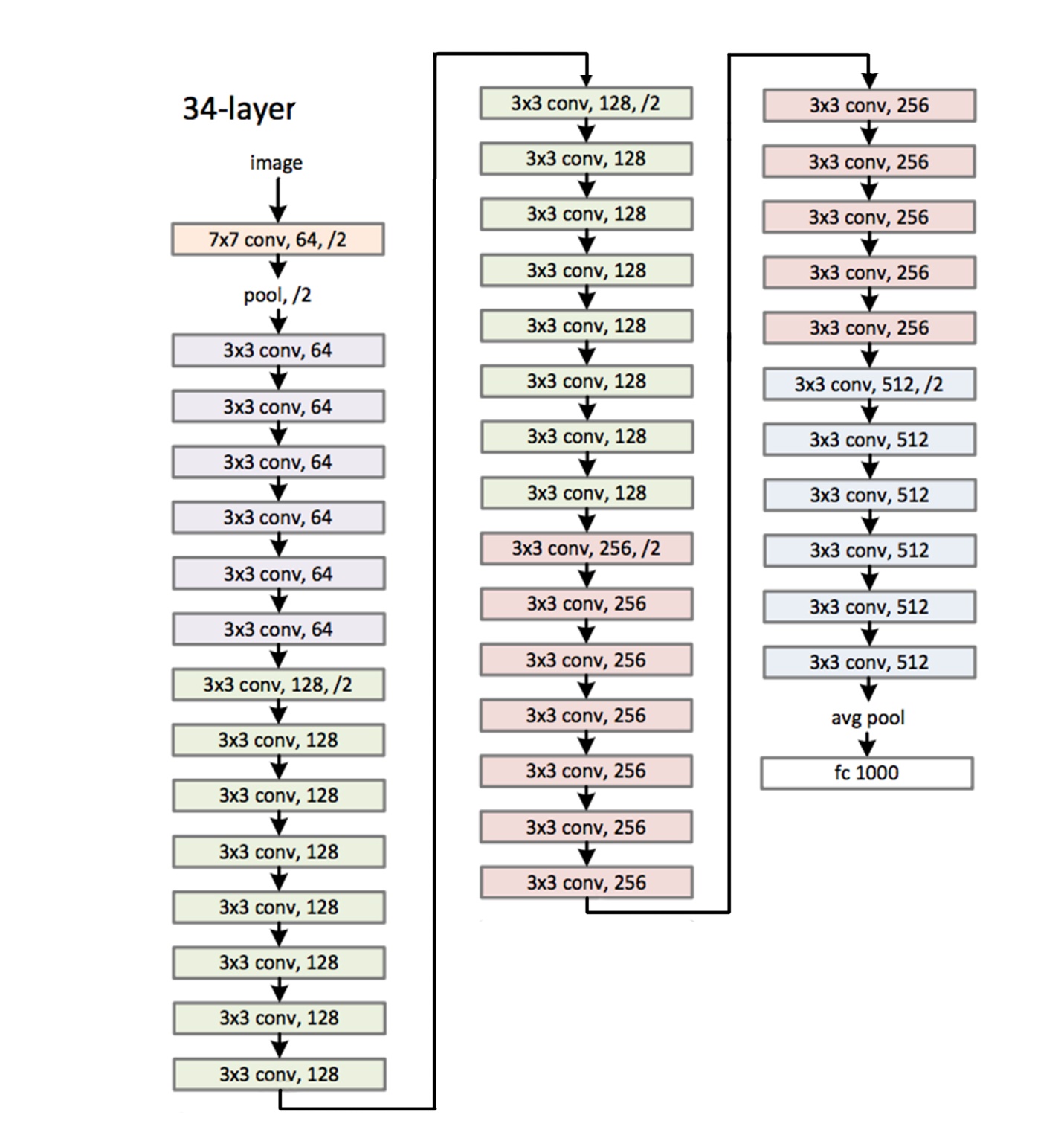


Рисунок 2. Архитектура НС «ResNet 34»

СНС «ResNet» состоит из 34 сверточных слоев. Отсутствие полносвязности сети, позволяют реализовать высокую глубину НС без риска переобучения. Тем не менее количество нейронов в НС «ResNet», а их в 34-х слойной сети более 20\*106, 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 использовались первоначальные веса нейронов полученные в результате обучения НС на задаче классификации изображений соревнования «ImageNet»[14]. Т.о. перед нами 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 [10, 11] для отображения пространства 1000-мерной размерности в 2-х мерное пространство. Результат для 600 изображений ЛА (300 F15 и 300 F16) приведен на рисунке Рисунок 3.

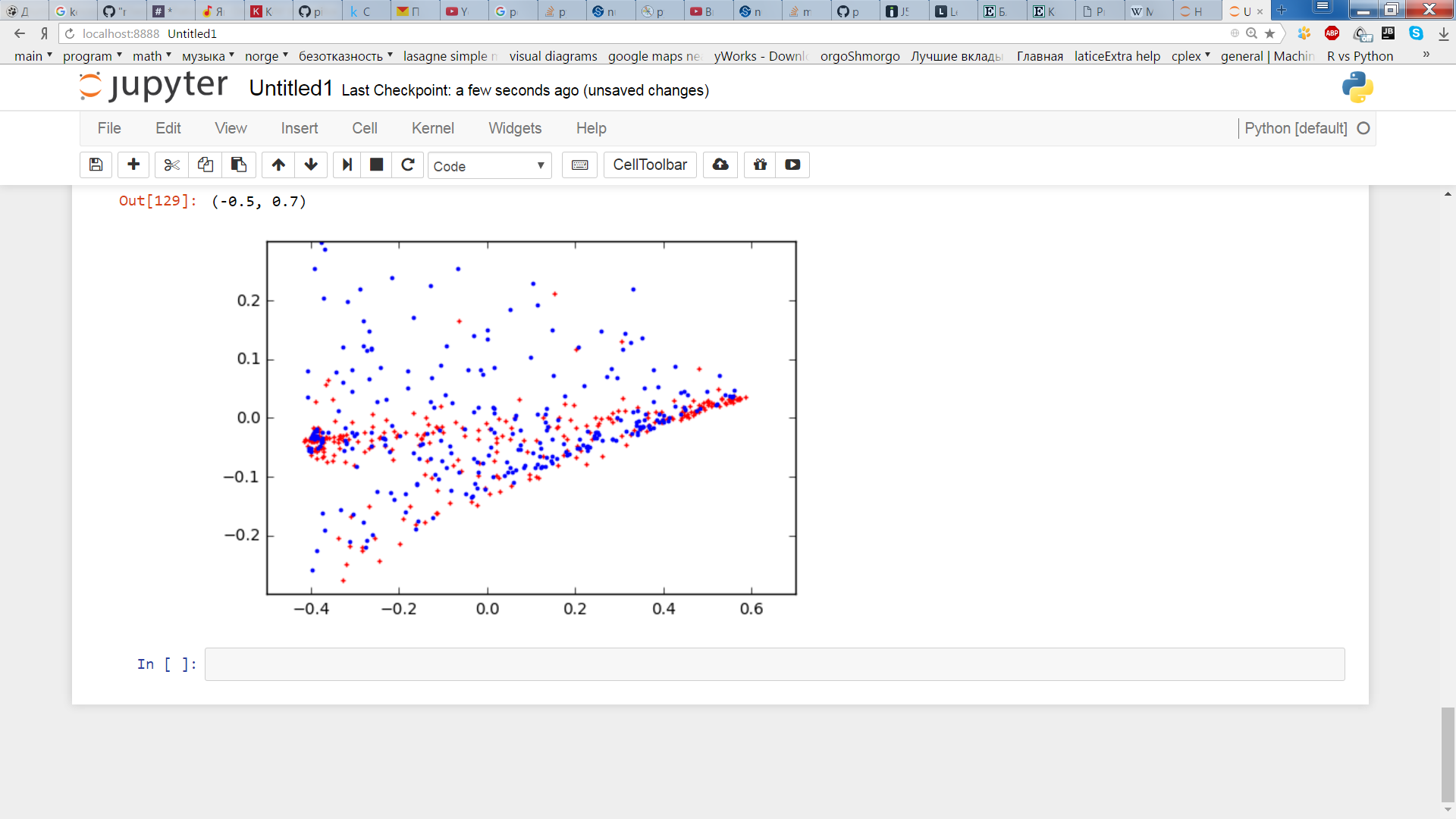


Рисунок 3. Проекция результатов выходного 1000-мерного слоя в 2-мерное пространство методом главных компонент. Красными ромбами обозначены выход для модели F16, синими точками – выход модели F15

После получения этого графика остается признать, что отличить модели F16 от F15 без некой доработки (адаптации) НС к задаче РГО типов ЛА 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). При адаптации был отрезан последний слой сети, и, вместо него, добавлен слой размерностью 2 (по числу исследованных в данной работе типов ЛА). Выход последнего слоя содержит вероятность принадлежности поданного на вход изображения соответствующему классу. Ответ выходного слоя (1,0) соответствует типу ЛА «F15» и (0,1) соответствует типу ЛА «F16». Процесс дообучения проходил в 2 этапа, на первом этапе были «заморожены» веса всех нейронов, кроме нейронов последнего, еще не обученного слоя. 80 эпох обучения позволили нам довольно точно найти значение нейронов, гарантирующее устойчивое решение задачи классификации, результаты оценки ошибки по эпохам приведены на рисунке Рисунок 4. Каждая эпоха включала в себя обучение по 1000 случайно выбранным изображениям. При этом из каждого изображения (которые имеют размер 300х300 пикселей) брался случайный участок 224х224 пикселя и изображение поворачивалось на случайный угол (0-150). Проверка (валидация) в каждой эпохе проводилась по 400 изображениям, которые так же выбирались случайно.

*Рисунок 4. Снижение оценки ошибки по эпохам обучения*

После обучения последнего слоя НС было выполнено дообучение всей сети целиком. Из-за высокой размерности задачи было выполнено всего 2 эпохи дообучения, каждая из которых включала в себя обучение на 10000 изображений и проверка (валидация) на 1000 изображений.

## Результаты численных исследований

После проведения дообучения и проведения валидации становится понятно, что НС получила способность различать сгенерированные искусственно изображения двух разных типов ЛА. Нас же интересует способность различать реальные изображения ЛА, для этого были взяты общедоступные фотографии F15, F16, типичные образцы которых приведены на рисунке Рисунок 5. Всего в проверке участвовало 20 изображений F15 и 21 изображение F16. Результаты численных исследований на реальных изображениях приведены в таблице **Ошибка! Источник ссылки не найден.**.

**



*Рисунок 5. Проверочные (реальные) изображения ЛА*

|  |  |  |
| --- | --- | --- |
| Тип ЛА | P(F15) | P(F16) |
| 20 изображений F15 | 0.801 | 0.199 |
| 21 изображение F16 | 0.222 | 0.778 |

Таблица 2. Результаты численных исследований на реальных изображениях. В таблице представлено среднее значение вероятности принадлежности изображений ЛА F15, F16 к классу F15, F16.

Отметим, что качество решения просело ухудшилось примерно в два раза, по сравнению с оценкой, полученной на сгенерированных (искусственных) изображениях.

Для проведения обучения использовался домашний компьютер в конфигурации: Процессор: core i74700, 8 гигабайт оперативной памяти, видеокарта: NVidia GeForce GT. Полный цикл обучения (80 эпох обучения последнего слоя и 2 эпохи обучения всей сети) занимали 40 часов. Применение новых быстрых видеокарт (для задач обучения рекомендуется использовать графические процессоры из-за их высокой разрядности) позволит снизить это время на порядок, к сожалению, при проведении исследований в распоряжении других ресурсов не было.

## Выводы

Подводя итоги отметим, что обе поставленные в данной работе цели достигнуты, СНС применимы в задачах классификации типов ЛА по изображению. Результаты, полученные в данной работе далеки от теоретического предела, который достижим в ближайшее время. Новейшие исследования СНС и прогресс в развитии вычислителей позволяет рассчитывать на решение этой задачи классификации в ближайшем бедующем с вероятностью не менее 95% по одной фотографии. Если представить себе, что наблюдение за ЛА продолжается некоторое время и мы имеем несколько изображений, полученных независимо (во времени, пространстве) то воспользовавшись теоремой о сложении вероятностей и пронормировав результат мы разом можем улучшить вероятность правильной классификации с 0.8 до 0.94.

Что же касается возможности применения сгенерированных данных для обучения, то и тут полученный результат можно считать удачным, обученная НС показывает приемлемый результат классификации реальных изображений (которые для обучения не использовались вовсе). В то же время необходимо отметить «проседание» результата при переходе к анализу реальных изображений. Для дальнего совершенствования НС представляется перспективным – наращивание данных для обучения (количество различных расцветов ЛА, типов фона, генерация изображений с различными типами боевой подвески и др.).

Стоит отметить что применяемая в данной работе СНС, так же, как и все самые лучшие СНС для задач классификации (Inception – GoogleLabs[714], VGG19 – Oxford Visual Geometry Group [15]) разработаны pа рубежом, в США и Великобритании.

## Ссылки

1. ResNet. Главная статья <https://arxiv.org/pdf/1512.00567v3.pdf>.
2. Модели F15, F16 https://github.com/piero10/AircraftsRecognition/tree/master/models
3. Реальные - проверочные изображений ЛА. <https://github.com/piero10/AircraftsRecognition/tree/master/images/realImages>
4. Хранилище исходного кода проекта. <https://github.com/piero10/AircraftsRecognition/tree/master>
5. Хранилище сгенерированных изображений ЛА. <https://github.com/piero10/AircraftsRecognition/tree/master/images/modelImages>
6. Библиотека низкого уровня для работы с НС - Theano. <http://deeplearning.net/software/theano/install_windows.html>
7. Библиотека высокого уровня для работы с НС - Keras. <https://keras.io/>
8. Соревнование по классификации изображений Imagenet. <http://image-net.org/challenges/LSVRC/2016/index>
9. Метод главных компонент. <http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>
10. Метод главных компонент. <http://setosa.io/ev/principal-component-analysis/>
11. Реализация .сверточной нейронной сети Resnet34 <https://github.com/KaimingHe/deep-residual-networks>
12. ResNet first place in classification task <http://image-net.org/challenges/LSVRC/2015/>
13. Сверточная НС <http://cs231n.github.io/convolutional-networks/>
14. GoogleLabs. Inception.<https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>
15. Oxford. Visual Geometry Group. <http://www.robots.ox.ac.uk/~vgg/>