

РАСПОЗНАВАНИЕ ДЕФЕКТОВ ДЛЯ ИНСПЕКЦИИ С УСТАНОВЛЕННОЙ НА КВАДКОПТОРЕ КАМЕРЫ

КЛЮЧЕВЫЕ СЛОВА fine-tuning, two-phase training, base classes, novel classes, few-shot, knowledge distillation, data augmentation, object detection, mean average precision

Объектом разработки является проведение экспериментов для сравнения перформанса разных моделей при ограниченном наборе данных

Целью работы является определение самой подходящей модели между DKAN, FSCE, Faster RCNN для применения ее на инспекцию дефектов поверхностей.

Методология проекта заключается в проведении следующих экспериментов:

1. Выбор модели для обучения на ограниченном наборе данных
Для определения самой подходящей модели, сравнивались модели Faster RCNN, DKAN, FSCE на разных настройках.
 - Faster RCNN обучение и оптимизация на базовых данных на 4 разных скоростях обучения.
 - DKAN 2-этапное обучение на целых данных.
 - FSCE 2-этапное обучение на целых данных.
 - верификация модели DKAN на дополненных данных
 - DKAN 2-этапное обучение на дополненных данных, и валидация на дополненных данных.

2. Верификация робастности выбранной модели из предыдущего эксперимента на разных преобразованиях и условиях имитирующих условий инспекции. Для этого, следующая аугментация применяются:

- размытие
- Высокая и низкая яркость
- повороты
- размытие, яркость и повороты все вместе

эта аугментация применяется в каждом эксперименте на трех сложностей.

3. предыдущий эксперимент повторяется после обучения модели на дополненных данных на размытие, яркость и повороты

Достижение цели проекта предполагает решение комплекса взаимосвязанных научно-технических задач:

- 1 провести обзор не менее 3 современных моделей и алгоритмов детектирования дефектов поверхностей по данным систем пассивного компьютерного зрения.
- 2 отдавать нейросетевым подходам, позволяющим осуществлять обучение и расширение классов детектируемых объектов при ограниченном наборе исходных данных.
- 3 Применить соответствующие преобразования к входным данным, чтобы имитировать вид с квадрокоптера.

Основными результатами являются:

1. DKAN предоставляет самую высокую точность из тестируемых моделей 0.683 mAP на обычных условиях и максимальная mAP равно 0.641 на условиях квадрокоптера с устойчивостью к освещением.
2. забивание базовых данных сходится при применении дистилляции знания
3. Модель чувствителен к вращению, и ее перформанс на больших углах вращения неприемлемый. В связи с этим рекомендуется собрать данных о дефектах с квадрокоптера из разных углов. Таким образом локализации дефектов будет более точно и соответственно модель станет инвариантной к поворотам.

Результаты проекта будут востребованы при выбора времени инспекции наблюдая условия погоды, ветра и яркости. Если условия инспекции выше чем приемлемые условия представлены в таблице то инспекция не будет проведена.

Содержание

ВВЕДЕНИЕ	6
1 Выбор направления исследований	7
2 Результаты теоретических и экспериментальных исследований	8
2.1 Влияние различных условий на перформанс детектора	8
2.1.1 Введение	8
2.1.2 Постановка задачи	8
2.1.3 Характеристика предлагаемого подхода	10
2.1.4 Достигнутые результаты	11
2.1.5 Оценка соответствия требованиям задач исследования	16
ЗАКЛЮЧЕНИЕ	17
Список использованных источников	18

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В настоящем отчете применяются следующие обозначения и сокращения:

catastrophic forgetting: Снижение производительности в некоторых классах после того, как модель обучается на новых различных классах.

fine-tuning: процедура обучения, в которой веса модели, обученной на определенном наборе данных, используются в качестве предварительных для обучения модели на другом наборе данных.

ограничивающие рамки: Прямоугольники, которые определяют местоположение объекта.

gt (ground truth): целевое значение выходного сигнала

DKAN: Dual knowledge align loss (сеть согласования двойных знаний)

ВВЕДЕНИЕ

Инспекция на рынке проводится с помощью квадрокоптера, управляемым человеком, появляется все больше возможностей сделать эту систему автономной, что можно сделать с помощью глубокого обучения, указав задачу как задачу обнаружения объекта. детекция дефектов страдает от нехватки данных, поскольку дефекты бывают редко, и постоянно появляются дефекты в новых видах. Поэтому важным свойством, которым должна обладать эта система, является способность постоянно обучаться при проведении инспекции и распознавать новые категории дефектов. Кроме того, эта система должна быть сложными мини-условия, с которыми может столкнуться квадрокоптер эти условия могут варьироваться от хорошего или рассеянного освещения до размытости из-за движения или различных углов обзора, поэтому при оценке этой модели необходимо учитывать эти условия

В число задач входит разработка:

- Модуль для применения геометрических преобразований, которые будут служить как увеличения данных так и имитация вида с квадрокоптера.
- модуль для выполнения обучения при настройке few-shot на двух этапах, base и fine-tuning.
- Модуль для применения разных моделей с data augmentation и без. и сравнения их выходы используя метрик mAP.
- Модуль для вычисления метрик mAP и создания confusion matrix

1 Выбор направления исследований

Обучение нераспределенной модели при ограниченном наборе исходных данных требует особых настроек, чтобы избежать overfitting. Это проблема называется дисбалансом данных. Один распространенный подход к этой проблеме называется fine-tuning. В этом подходе рассматривается обучение на двух этапах. В первом этапе (Base Training), модель обучают на категориях, обладающих большим количеством данных. Эти данные называются базовыми данными Base data. Во втором этапе, модель переобучают на новых категориях, неизвестных модели. Но если модель обучается на новых категориях, то модель забывает базовые данные. Эта проблема называется catastrophic forgetting. Соединение новых данных с базовыми данными и обучение модели с нуля требует дополнительное время и считается неэффективным процессом. В этом реферате, рассматриваются подходы, которые решают эти проблемы без дополнительных потерь. Целью является определение подходящей глубокой модели для детекции дефектов

2 Результаты теоретических и экспериментальных исследований

2.1 Влияние различных условий на перформанс детектора

2.1.1 Введение

Для того чтобы запустить детектор на квадрокоптере, необходимо чтобы он владел несколькими особенностями. Детектор должен быть способен распознавать дефектов не смотря на текущие условия погоды. Эти условия могут варьироваться в связи с обстановкой инспекции. Для того чтобы убедить что перформанс модели будет высокий в таких условиях нужно эмитировать эти условия при обучении и валидации. Для этого экспериментов должен быть проведен чтобы исследовать влияние разных условий на перформанс модели.

2.1.2 Постановка задачи

Датасет состоит из 6 классов. В каждом из них есть 300 экземпляров. Модели обучаются на 240 экземпляров, и тестируются на 60. На первом этапе обучение 240 экземпляров используются. А на втором только 5 из каждого класса. На рисунке показана распределение экземпляров на втором этапе относительно целого распределения данных.

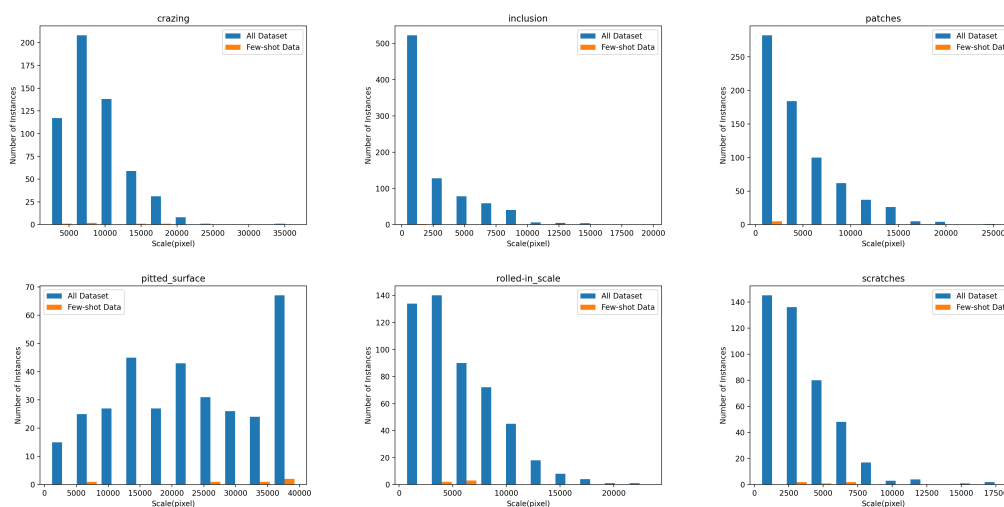


Рисунок 1 — — используемое распределение датасета NEU DET на втором этапе. Оранжевый определяет shots, а синий определяет целое распределение. С рисунком видно, что обобщение 5 эк на целое распределение является сложной задачей

Обучение проведется на следующий пайплайн

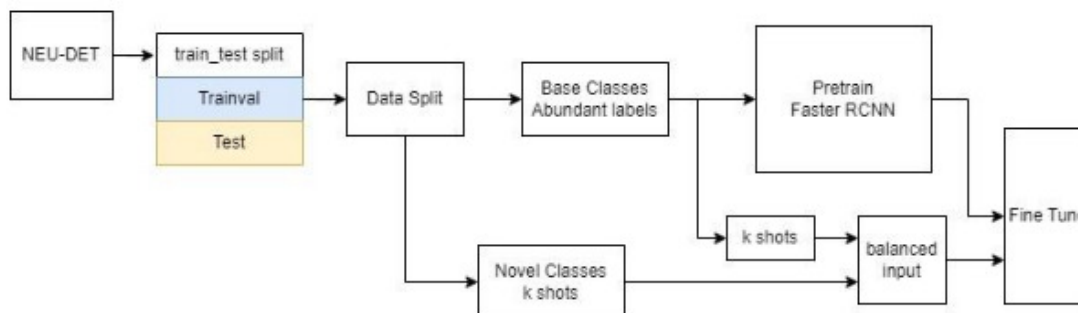


Рисунок 2 — — Пайплайн обучения

входные данные разделятся на train и test. На первом этапе, используются все данные базовых классов, и входят в модель Faster RCNN. после того, как модель дообучается, ее веса переходят во

второй этап. из данных train произвольно берутся K shots из базовых и новых данных.

2.1.3 Характеристика предлагаемого подхода

На первом этапе обучается Faster RCNN. на втором этапе обучения DKAN копируется модель Faster RCNN два раза. И переходят две потери с учителя на студента.

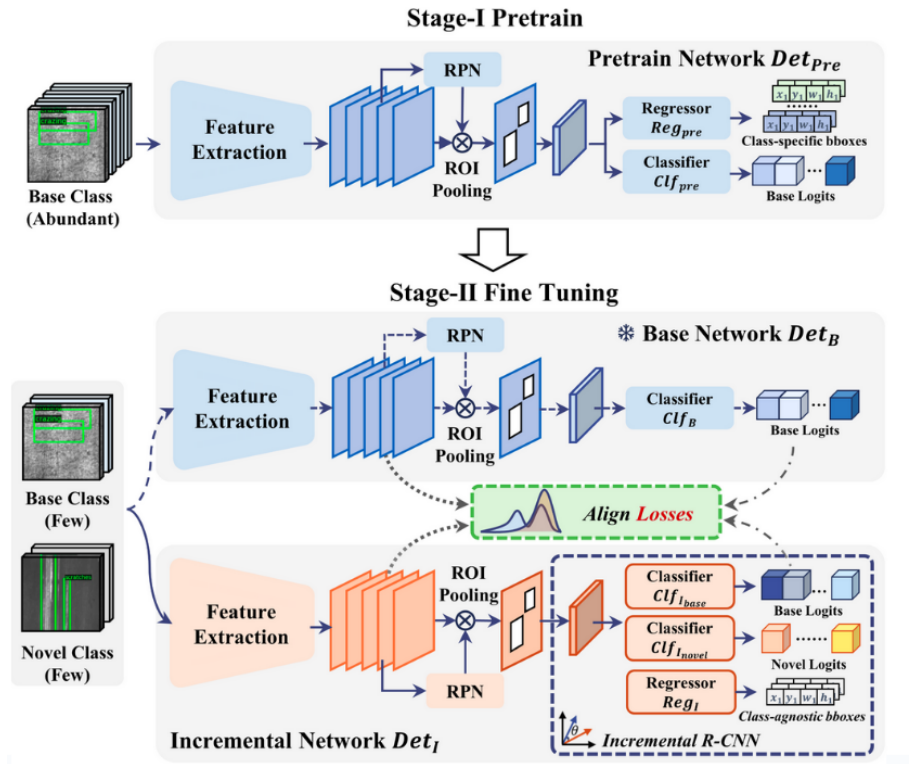


Рисунок 3 — — Архитектура модели DKAN

потери вычисляются следующим образом.

$$Loss_{FKA} = \sum_{i=2}^5 \varphi(p_i^T, p_i^S)$$

$$L_{LKA} = \frac{\tau^2}{C_{base}} \sum_{i=1}^{C_{base}} \phi(n_{S_{base}}^i) \log \left(\frac{\phi(n_{S_{base}}^i)}{\phi(n_r^i)} \right)$$

Где FKA сокращение Feature Knowledge Align loss, а LKA сокращение Logit Knowledge Align loss.

2.1.4 Достигнутые результаты

Экперемент 1

Дополненные данные были использованы для имитации вида с квадранта. Вниз предоставлены экземпляры оригинальных данных. Здесь аугментация ограничена на вращательное преобразование в углу не больше 15 градусов и угол shear равен 15 градусов.

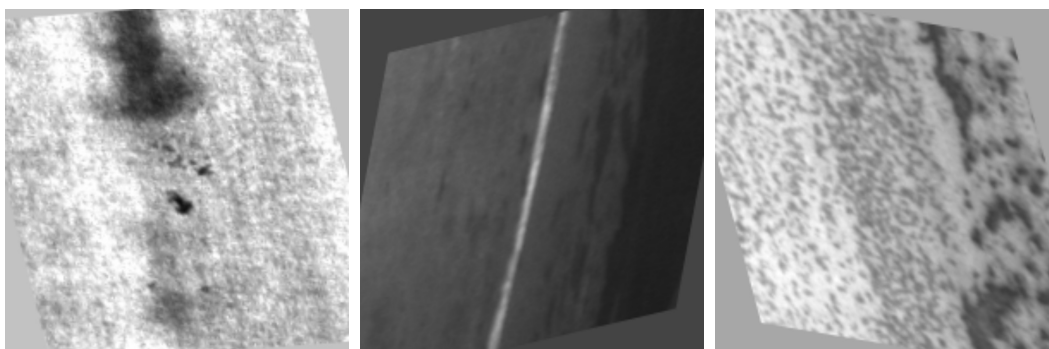


Рисунок 5 — — аугментация используемая в первом эксперименте

Оптимизация Faster RCNN на первый этап. Для этого, были использованы разные скорости обучения

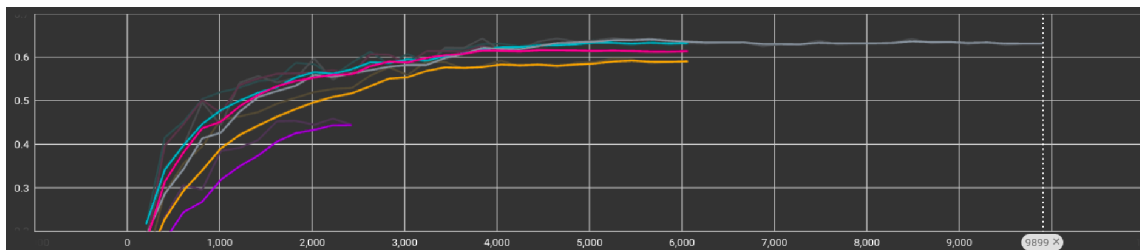


Рисунок 6 — — Оптимизация Base Training

скорости обучения 0.01 и 0.02 показывают похожую динамику. Поэтому был выбран 0.01

После обучения DKAN на трех скоростях обучения, была выбрана скорость обучения 0.02

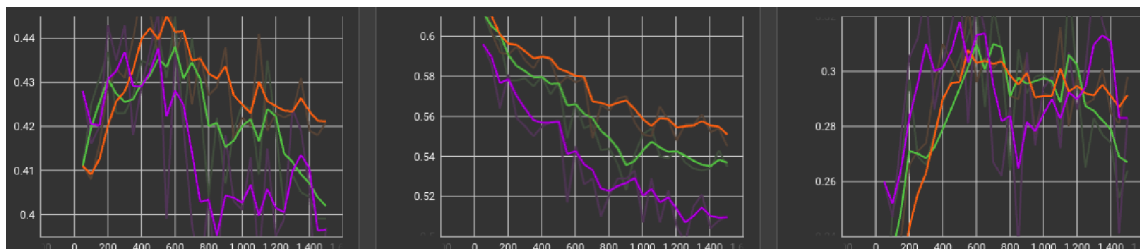


Рисунок 7 — — Оптимизация fine-tuning модели DKAN

при обучении и тестировании учитываются различные настройки. эти настройки сравнивают производительность модели без каких-либо дополнений и ее производительность на дополненных тестовых данных, затем проводится эксперимент, в котором модель обучается на дополненных данных, а затем тестируются дополненные данные

	train aug	test aug	mAP Base	mAP Novel	mAP ALL
Faster RCNN (240 shots)	-	-	0.644	-	-
DKAN (5 shots)	-	-	0.558	0.334	0.446
	x	-	0.569	0.346	0.457
	-	x	0.129	0.229	0.179
	x	x	0.148	0.239	0.193
FSCE (5 shots)	x	x	0.13	0.189	0.16

Таблица 1 — mAP обученных моделей на базовых и новых данных

Производительность модели значительно снижается после ее тестирования на дополненных данных. Однако после применения аугментации данных во время обучения, mAP на тестовых данных увеличивается.

при тестировании моделей на дополненных данных, DKAN предоставляет самую высокую точность на базовых и новых классах

Эксперимент 2

В этом эксперименте дополненные данные были премены с 3 различными сложностями освящении, размытия и вращении, примененных отдельно и вместе. освящение варьируется в интервале $[-0.15, 0.15]$ на легкие настройки, в интервале $[-0.3, 0.3]$ на настройки средно сложности, и в интервале $[-0.45, 0.45]$ на сложные настройки. размытие варьируется в интервалах $[5, 10]$, $[10, 20]$, $[15, 30]$ последовательно. вращения $[-30, 30]$, $[-60, 60]$, $[-90, 90]$ последовательно.

на следующем рисунке представлены дополненные данные с тремя различными дополнениями одновременно для трех различных трудностей

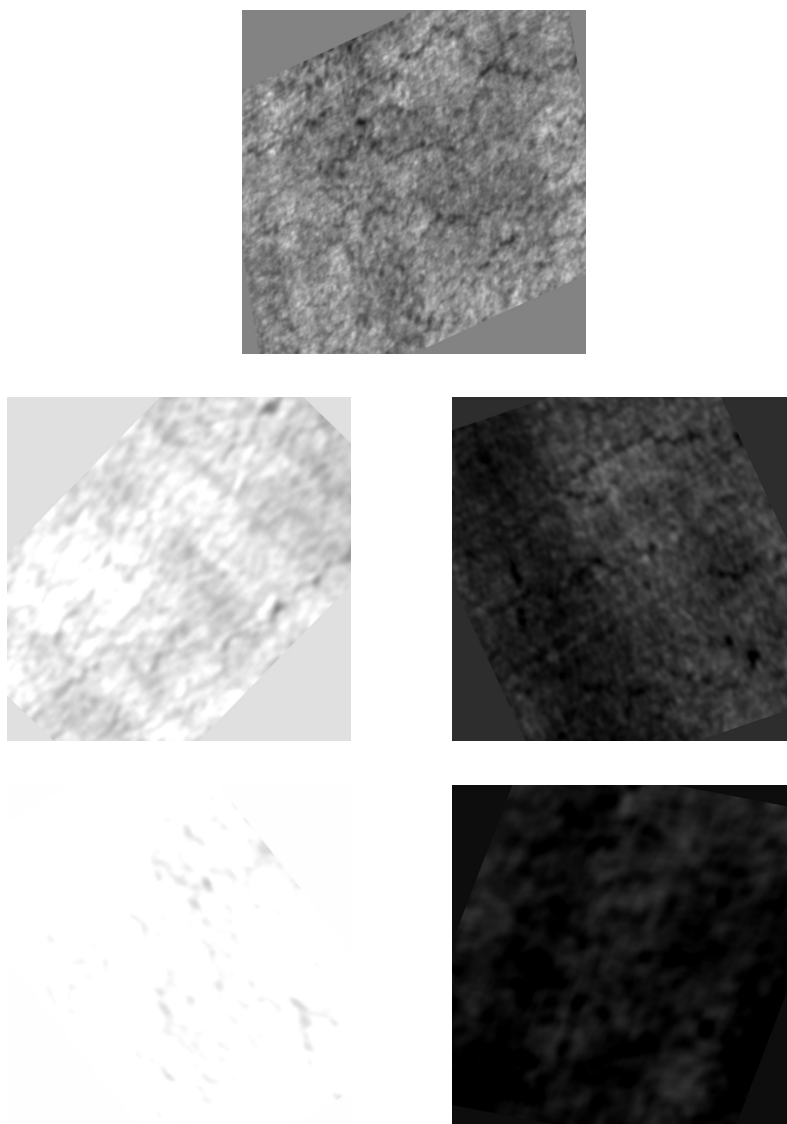


Рисунок 9 — — дополненные данные в первой строке показано изображение с легкой настройкой, во второй строке показаны два изображения со средней сложностью, в третьей строке показаны два изображения с высокой сложностью. Изображения паказоны на самой большой виличиной в интервале

Модель DKAN была обучена на всех данных train из всех категорий. Перед применением преобразования, mAP модели равно 0.683.

Augmentation	mAP		
	Easy 30%	Medium 60%	Hard 90%
blur	0.616	0.435	0.331
rotate	0.257	0.169	0.143
bright	0.641	0.568	0.5
rotate,blur,bright	0.216	0.084	0.036

Таблица 2 — — Результаты модели DKAN при применении разных преобразованиях на разных сложностях

Эксперимент 3

В этом эксперименте DKAN Обучается на дополненных данных где применяются 3 аугментации вместе. mAP модели когда тестируется на недополненных данных равно 0.543

Augmentation	mAP		
	Easy 30%	Medium 60%	Hard 90%
blur	0.582	0.606	0.598
rotate	0.241	0.162	0.146
bright	0.542	0.529	0.467
rotate,blur,bright	0.264	0.202	0.146

Таблица 3 — — Результаты модели DKAN, которая предварительно обучена на 3 аугментациях вместе

На этих настройках Модель стала более робастным к размытости. Это было заметно когда перформанс модели не изменился после применения отмытия на разных сложностях.

По сравнению с предыдущим экспериментом перформанс поднялся только в случае когда все аугментации были применены вместе.

2.1.5 Оценка соответствия требованиям задач исследования

Предлагаемая модель предоставляет mAP 0.68 При обучении на 30 экземплярах. для детекции дефектов необходимо иметь не меньше 5 экземпляров каждого дефекта, но при этом потребуется вторая проверка чтобы убедиться что модель не пропустила дефекты. Для того чтобы обеспечить высокий уровень доверенности, необходимо иметь не менее экземпляров. Модель готова функционировать в разных условиях освещения, при размытости возникающей при перемещении квадрокоптера. Однако она не может распознать дефекты когда смотрятся из разных углов а именно могут распознать их если подкоптер находится прямо вверх. Для решения этой проблемы необходимо собрать данные квадрокоптера из разных углов. Модель может изучать постоянно новые типы дефектов и научиться распознавать их.

ЗАКЛЮЧЕНИЕ

По результатам выполнения проекта были получены следующие научно-технические результаты:

1. Модель получает высокую точность на зафиксированном положении. Модель относительно устойчивая к освещению. Но страдает от большой неробастности против вращений, и размытия
2. Обучение моделей на разных агмутациях увеличивает устойчивость к размытию, но точность в целом снижается. В этом случае перформанс увеличивается только при применении разных условиях вместе.
3. Модель не может обобщаться и не может функционировать на квадрокоптере, если он не обучён на таких же условиях, исключая освящения. Доказательством того является снижение точности при тестировании моделей на дополнительных данных не обучая её предварительно на них.
4. Дистилляции знаний, замедляет забывание базовых данных, и перформанс модели на них сходится.

Таким образом, все задачи, поставленные для проекта, успешно выполнены.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Chen, Li, Chunyan Yu, and Lvcai Chen. “A New Knowledge Distillation for Incremental Object Detection.” In 2019 International Joint Conference on Neural Networks (IJCNN), 1–7, 2019.
<https://doi.org/10.1109/IJCNN.2019.8851980>.
2. Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation.” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, November 11, 2013.
<https://doi.org/10.1109/CVPR.2014.81>.
3. Lin, Tsung-Yi, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. “Feature Pyramid Networks for Object Detection.” In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 936–44. Honolulu, HI: IEEE, 2017. <https://doi.org/10.1109/CVPR.2017.106>.
4. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” In Advances in Neural Information Processing Systems, Vol. 28. Curran Associates, Inc., 2015.

5. Sun, Bo, Banghuai Li, Shengcai Cai, Ye Yuan, and Chi Zhang.
“FSCE: Few-Shot Object Detection via Contrastive Proposal Encoding.”
2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition
(CVPR), June 2021, 7348–58.
<https://doi.org/10.1109/CVPR46437.2021.00727>.