УЛУЧШЕННЫЙ АЛГОРИТМ ИНТЕРПОЛЯЦИИ ДЛЯ НЕЙРОННЫХ СЕТЕЙ, ИСПОЛЬЗУЕМЫХ ДЛЯ РАСПОЗНАВАНИЯ НЕБОЛЬШИХ ОБЪЕКТОВ

IMPROVEMENT OF INTERPOLATION ALGORITHM FOR NEURAL NETWORKS USED TO IDENTIFY SMALL-SIZE OBJECTS

УДК 339.54.012+338.001.36

Лоу Гуаньпин, магистрант,
Московский государственный технический университет, имени Н.Э.Баумана.

Лоу Гуаньпин aliksandal@gmail.com

Аннотация

В последние несколько десятилетий совершенствование различных теоретических основ привело к очень быстрому развитию в области систем машинного зрения, и многие нейронные сети для распознавания имеют очень стабильную точность распознавания в специальных сценариях, но методы мониторинга в более сложных средах все еще нуждаются в дальнейшем совершенствовании. Проблема неточного или даже необнаружения, вызванная слишком маленькими пикселями распознаваемых целей на изображениях, представляет собой большую проблему для дальнейших исследований в области распознавания и отслеживания целей в реальных сложных средах.

Для того чтобы повысить точность распознавания малоразмерных целей, многие ученые предложили много очень эффективных методов, таких как использование многомасштабного обнаружения карты признаков[1], получение большего количества воспринимаемых полей с помощью Dilated Convolution[2]

и обучение с помощью многомасштабных образцов[3] и т.д. Однако в некоторых сценариях, таких как изображения, полученные БПЛА с воздуха, объекты распознавания в этих сценариях обычно небольшие, и поэтому эти алгоритмы не обладают хорошей точностью распознавания В данной работе мы предлагаем метод интерполяции, который увеличивает изображение и сохраняет соответствующую четкость, а затем улучшает распознавание небольших объектов путем обработки изображения по частям после увеличения размера. Этот метод можно комбинировать с ранее упомянутыми алгоритмами для дальнейшего улучшения распознавания небольших объектов.

Annotation

In the last few decades, improvements in various theoretical frameworks have led to very rapid developments in the field of machine vision systems, and many neural networks for recognition have very stable recognition accuracy in ad hoc scenarios, but monitoring methods in more complex environments still need further improvement. The problem of inaccurate or even non-detection caused by too small pixels of recognized targets in images is a big challenge for further research in target recognition and tracking in real-world complex environments.

In order to improve the recognition accuracy of small-scale targets, many scientists have proposed many very effective methods, such as using multiscale feature map detection, obtaining more perceived fields with Dilated Convolution and learning with multiscale samples, etc. However, in some scenarios, such as aerial drone images, the recognition objects in these scenarios are usually small, and therefore these algorithms do not have good recognition accuracy In this paper, we propose an interpolation method that enlarges the image and maintains appropriate sharpness, and then improves recognition of small objects by processing the image piece by piece after the size increase. This method can be combined with the previously mentioned algorithms to further improve the recognition of small objects.

Ключевые слова: глубокое обучение, нейронные сети, системы компьютерного зрения, распознавание объектов, алгоритм интерполяции.

Keywords: deep learning, neural networks, computer vision systems, object recognition, interpolation algorithm.

В данной работе в качестве распознающей нейронной сети выбрана сеть YOLO-V3, которая будет обработана несколькими популярными алгоритмами интерполяции и проанализирована на основе четкости алгоритма интерполяции после обработки, предложено улучшение алгоритма интерполяции и проведено сравнение с популярными алгоритмами интерполяции. Наконец, обработанные изображения применяются к нейронной сети YOLO-V3[4] для распознавания целей, и результаты распознавания сравниваются.

Алгоритм интерполяции и чёткость изображение

В процессе геометрического преобразования изображения невозможно присвоить значения некоторым точкам пикселей напрямую; например, увеличение изображения в два раза неизбежно приведет к появлению дополнительных точек пикселей, которые не могут быть отображены напрямую, и их значения определяются путем интерполяции. Таким образом, возникает алгоритм интерполяции изображения. Популярными алгоритмами интерполяции являются интерполяция ближайших соседей, билинейная интерполяция и бикубическая интерполяция[5].

Интерполяция ближайших соседей: Это самый простой метод интерполяции, не требующий вычислений, при этом искомому пикселю присваивается градация серого цвета ближайшего соседа искомого пикселя из четырех соседних пикселей искомого пикселя.

Билинейная интерполяция: Основная идея этого метода интерполяции заключается в том, чтобы найти значения пикселей на основе четырех соседних значений пикселей точки после трех вычислений линейной интерполяции. Это делается путем выполнения двух вычислений линейной интерполяции в направлении X, а затем одного вычисления интерполяции в направлении Y.

Бикубическая интерполяция: При таком методе значение функции f в точке (x,y) может быть получено взвешенным средним из шестнадцати ближайших выборочных точек прямоугольной сетки, где необходимо использовать две полиномиально интерполированные кубические функции, по одной в каждом направлении.

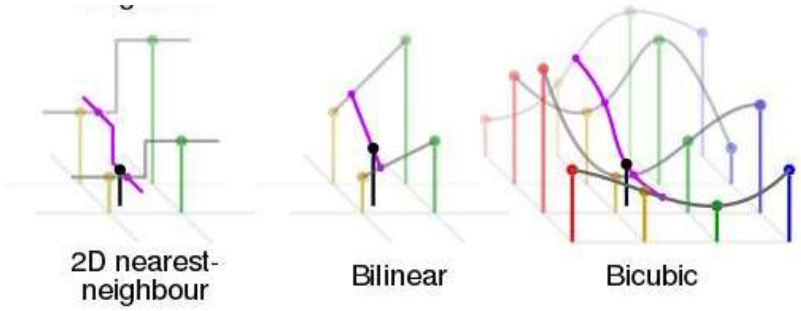


Рис 1. Популярные алгоритмы интерполяции

Эти алгоритмы интерполяции часто имеют хорошую производительность, когда увеличение изображения не слишком велико. Однако, как говорилось выше, на изображениях, полученных БПЛА с воздуха, объекты распознавания маленькие, и если увеличение составляет менее двух раз, это не имеет существенного значения для улучшения результатов распознавания, поэтому часто необходимо увеличивать изображения в 3-4 раза. Поэтому особенно важно сохранить признаки объекта распознавания после многократного увеличения изображения[6].

Мы используем алгоритм интерполяции, который напрямую связан с исходными пикселями изображения. Он использует метод, похожий на сверточные фильтр, для расширения значений пикселей, что делает изображение более четким и обеспечивает связь пикселя на изображении с соседними пикселями, благодаря чему лучше сохраняются признаки объекта.

Этот алгоритм реализуется следующим образом:

Допустим, что A - матрица изображения, а B - фильтр алгоритма интерполяции, если мы хотим увеличить изображение в 4 раза, то размер матрицы B равен 4*4. Построив матрицу B и перемножив ее с матрицей изображения A, получим увеличенную матрицу изображения[7].

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix}$$

Затем для каждого пикселя $a_{ij}(i = 1, 2, \dots, N; j = 1, 2, \dots, M)$ в матрице изображения A с разрешением N * M постройте интерполяционный фильтр. Сначала пусть b_{11} = 1;

$$A_{i,j} = \begin{bmatrix} a_{i,j} & a_{i,j+1} & a_{i,j+2} & a_{i,j+3} \\ a_{i+1,j} & a_{i+1,j+1} & a_{i+1,j+2} & a_{i+1,j+3} \\ a_{i+2,j} & a_{i+2,j+1} & a_{i+2,j+2} & a_{i+2,j+3} \\ a_{i+3,j} & a_{i+3,j+1} & a_{i+3,j+2} & a_{i+3,j+3} \end{bmatrix}$$

Затем сравним элементы $a_{u,v}(u \neq i\ , v \neq j)$ и $a_{i,j}$ в матрице $A_{i,j}$. Если $a_{u,v}$ < $a_{i,j}$, то элемент $b_{m,n}$ в матрице фильтра, соответствующий позиции $a_{u,v}$, равен$b_{m,n} = 1 - w * (\frac{a_{i,j} - a_{u,v}}{a_{i,j}} + rand(1))$ где: rand (1) в формуле - это случайная величина функции распределения Гаусса на интервале [0,1],w - константа,в данной работе она принимает значение 0,01. Если $a_{u,v} > a_{i,j}$, то элемент$b_{m,n} = 1 + w * (\frac{a_{u,v} - a_{i,j}}{a_{u,v}} + rand(1))$; Если $a_{u,v} = a_{i,j}$, то элемент $b_{m,n} = 1$;

Это увеличивает длину и ширину исходного изображения в 4 раза, но в некоторой степени сохраняет резкость изображения. Увеличение изображения A может быть достигнуто путем построения серии матриц фильтров B с использованием вышеописанного метода построения и последующего поочередного умножения их на элементы изображения A.

В конкретном случае, пусть матрица A обозначает матрицу пикселей части изображения следующим образом.

$$A = \begin{bmatrix} 130 & 164 & 88 & 182 \\ 173 & 116 & 171 & 147 \\ 101 & 96 & 138 & 155 \\ 67 & 51 & 43 & 99 \end{bmatrix}$$

где значения пикселей от 0 до 255. Затем, используя 125 в качестве $a_{i,j}$, постройте матрицу B, как показано ниже.

$$B = \begin{bmatrix} 1 & 1.04 & 0.97 & 1.06 \\ 1.05 & 0.98 & 1.04 & 1.02 \\ 0.98 & 0.97 & 1.01 & 1.03 \\ 0.95 & 0.96 & 0.96 & 0.98 \end{bmatrix}$$

Умножьте полученную в этой точке матрицу B на значения пикселей $a_{i,j}$, чтобы получить матрицу интерполяции;

$$a_{i,j} * B = \begin{bmatrix} 130 & 135 & 126 & 138 \\ 137 & 127 & 135 & 133 \\ 127 & 126 & 131 & 134 \\ 124 & 125 & 125 & 127 \end{bmatrix}$$

Применяя расширение пикселей к каждому пикселю исходного изображения таким образом, можно получить более четкое изображение.

Результат

Когда мы получаем увеличенное изображение, поскольку размер изображения увеличивается в четыре раза, если мы используем обнаружение сети напрямую, это может привести к медленному обнаружению или не получить хорошие признаки объектов, поэтому мы разделяем увеличенное изображение на 16 блоков для распознавания отдельно. YOLO-V3 выбран для нейронной сети распознавания.



Рис 2. Исходное изображения

Мы использовали исходное изображение для входа в нейронную сеть YOLO-V3 и получили следующие результаты:

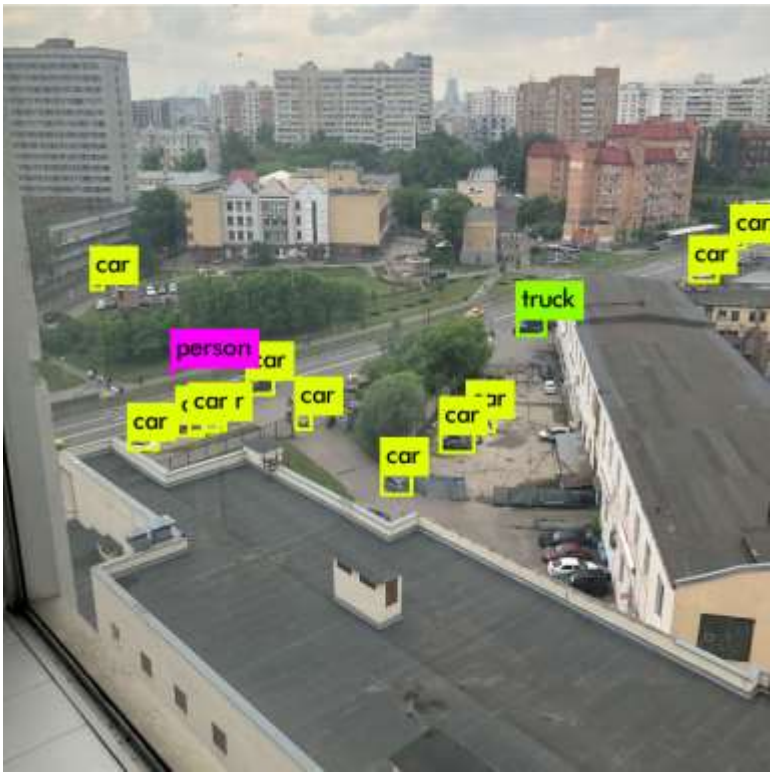
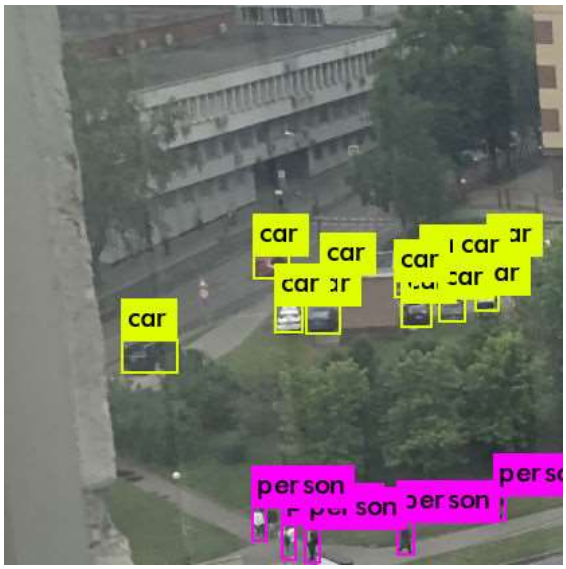
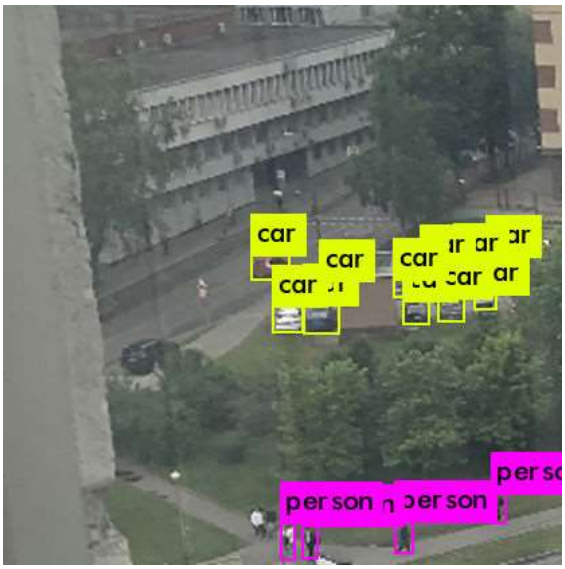


Рис 3. Результат распознавания исходного изображения

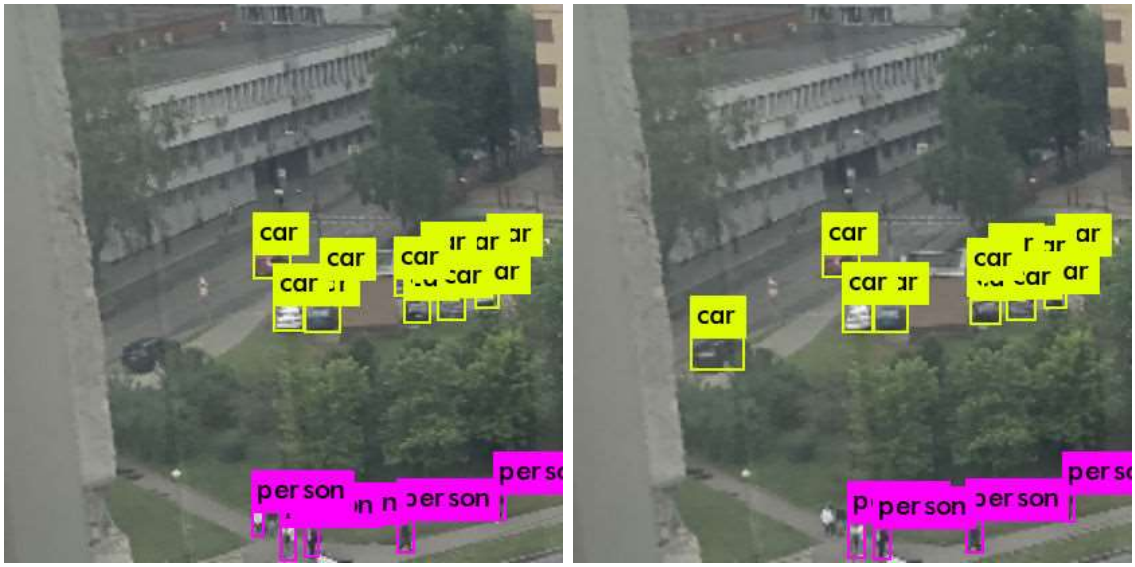
Из результатов видно, что когда размер изображения не увеличен, многие объекты не распознаются из-за их маленького размера. Далее мы поделили изображение на четыре части по длине и ширине, последовательно увеличили масштаб, а затем ввели нейронную сеть YOLO-V3 и получили следующие результаты:



(a)



(b)



(c) (d)

Рис 4. (a) результат интерполяции с сверточным фильтром (b) результат бикубической интерполяции (c) результат билинейной интерполяции (d) результат интерполяции ближайших соседей

Сравнение результатов показывает, что больше информации об объекте может быть получено из увеличенного изображения алгоритмом интерполяции и, следовательно, больше объектов может быть распознано, в то время как наш алгоритм интерполяции способен сохранить больше информации о признаках и, следовательно, распознавание более эффективно.

Заключение

По результатам обнаружения YOLO-V3, наш алгоритм интерполяции правильно распознает всех людей и автомобили на изображениях, имеет более высокую точность распознавания, чем другие алгоритмы интерполяции, и не пропускает ни одной объекта, но если он будет использоваться в реальных сценариях, необходимо найти компромисс между скоростью обработки компьютера и точностью распознавания.

Список литературы

1. Liu W. et al. Ssd: Single shot multibox detector //European conference on computer vision. – Springer, Cham, 2016. – С. 21-37.

2. Li Y., Zhang X., Chen D. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 1091-1100.
3. Lin T. Y. et al. Feature pyramid networks for object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 2117-2125.
4. Redmon J. et al. You only look once: Unified, real-time object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 779-788.
5. Rahim R. et al. Comparison Searching Process of Linear, Binary and Interpolation Algorithm //Journal of Physics: Conference Series. – IOP Publishing, 2017. – Т. 930. – №. 1. – С. 012007.
6. Perrone D., Favaro P. A clearer picture of blind deconvolution //arXiv preprint arXiv:1412.0251. – 2014.
7. Schacke K. On the kronecker product //Master's thesis, University of Waterloo. – 2004.