

УДК 629.5.073.4

DOI: 10.34046/aumsuomt96/12

РЕШЕНИЕ ПРАКТИЧЕСКОЙ ЗАДАЧИ РАСПОЗНАВАНИЯ ЛЕДОВЫХ ТРЕКОВ И ОПРЕДЕЛЕНИЯ ИХ НАПРАВЛЕНИЯ С ПРИМЕНЕНИЕМ МАШИННОГО ЗРЕНИЯ И ОБУЧЕНИЕМ НЕЙРОННОЙ СЕТИ

В.В. Тульчинский, аспирант

А.Л. Боран-Кешишьян, кандидат технических наук

В статье рассмотрена возможность использования машинного зрения при ледовом плавании с целью автоматизации процесса поиска «треков» в ледяных полях. Представлена технология обработки изображений и детектирования границ ледового разлома с применением сверточной нейронной сети архитектуры MobileNet и глубокого обучения. Описан способ поиска теоретической линии ледового «трека» с помощью линейной регрессии. Создано программное средство, способное производить автоматизированный поиск ледовых «треков» и производить расчет их отклонения от курса судна. Рассмотрены ограничения и недостатки программы, а также ее практическая значимость в рамках перспектив внедрения автоматических систем судовождения на примере СПГ газовоза типа Arc7.

Ключевые слова: машинное зрение, арктическая навигация, навигация во льдах, автоматизация судовождения, ледовый «трек», сверточная нейронная сеть, глубокое обучение, нейросети, линейная регрессия, СПГ газовоз, Arc7 газовоз.

The article studies the possibility of using the machine vision while ice navigation in order to automate the process of detecting "tracks" in ice fields. The technology of image processing for detecting ice fractures' boundaries with help of MobileNet convolutional neural network architecture and deep learning is presented. A method of calculating the theoretical line of the ice "track" using linear regression is described. A software product has been created and it is capable to perform ice "tracks" automated detection and to calculate their deviation from the true course of the vessel. The limitations and disadvantages of the program, as well as its practical importance in the prospects of automatic navigational systems, are described. Program shows an example of implementation on Arc7 LNG carrier.

Keywords: machine vision, arctic navigation, ice navigation, automation of navigation, "tracks" in ice, convolutional neural network, deep learning, neural networks, linear regression, LNG vessel, Arc7 LNG.

Введение. При ледовой навигации наиболее удобным инструментом для преодоления ледовых полей является поиск ледовых разломов и ледовых «треков», оставленных после прохождения других судов. Их использование для следования собственного судна позволяет комплексно повысить уровень безопасности ледового плавания, а также его общую экономическую эффективность. Современные перспективы развития судовых навигационных систем тесно связаны с внедрением автоматизированных процессов. А автоматизированные процессы ориентации в пространстве и навигации сложно представить без современных систем машинного зрения.

Как было рассмотрено в статье [5] одной из приоритетных задач автоматизации ледовой проводки судна, является автоматический поиск ледовых разломов и «треков» с применением технологии машинного зрения. Было выяснено, что в рамках поставленной задачи определения азимутального направления линии ледового «трека», можно использовать несколько подходов, а именно:

- Определение с помощью детектора границ. Используется дифференциальный подход (оператор Собела, алгоритм Канни, преобразование Хафа).

- Семантическая сегментация «областей интереса» кадров, выделение фрагментов. Обучение глубинной нейросети на одной из известных моделей.

Особенность «сцены» ледовых полей такова, что существует достаточное количество объектов, которые можно классифицировать по-разному. Это и различные виды ледовой крошки, и торосы, и снег на ледовых полях, и различные типы ледовых разливов. Формулировка задачи состоит в поиске ледового «трека», то есть множества объектов с определённой классификацией. Таким образом, было определено, что удовлетворить условиям задачи может только подход, заключающийся в использовании и обучении глубинной нейросети [5].

В рамках решения поставленной задачи, был произведен сбор видеоматериалов на борту СПГ газовоза типа Arc7 во время ледовой навигации.

Различные фрагменты видеоряда были записаны со стационарной камеры, расположенной на левом крыле судна. Камера была расположена таким образом, что ее главная оптическая ось лежит в плоскости левого борта судна, которая в свою очередь, параллельна ДП. Это позволило произвести ориентацию главной оптической оси по курсу судна, что упростило дальнейшие вычисления. При этом, камера установлена так, что ее фокальная плоскость параллельна плоскости

мидель-шпангоута судна. Таким образом, полученные изображения из видеоряда условно находятся в плоскости, перпендикулярной ОП, или плоскости водной поверхности на рассматриваемом участке акватории (Рисунок 1). При этом

фрагменты видеоряда были записаны в дневное время, но при различной освещенности сцены. На данных фрагментах запечатлены различные структуры ледовых «треков».

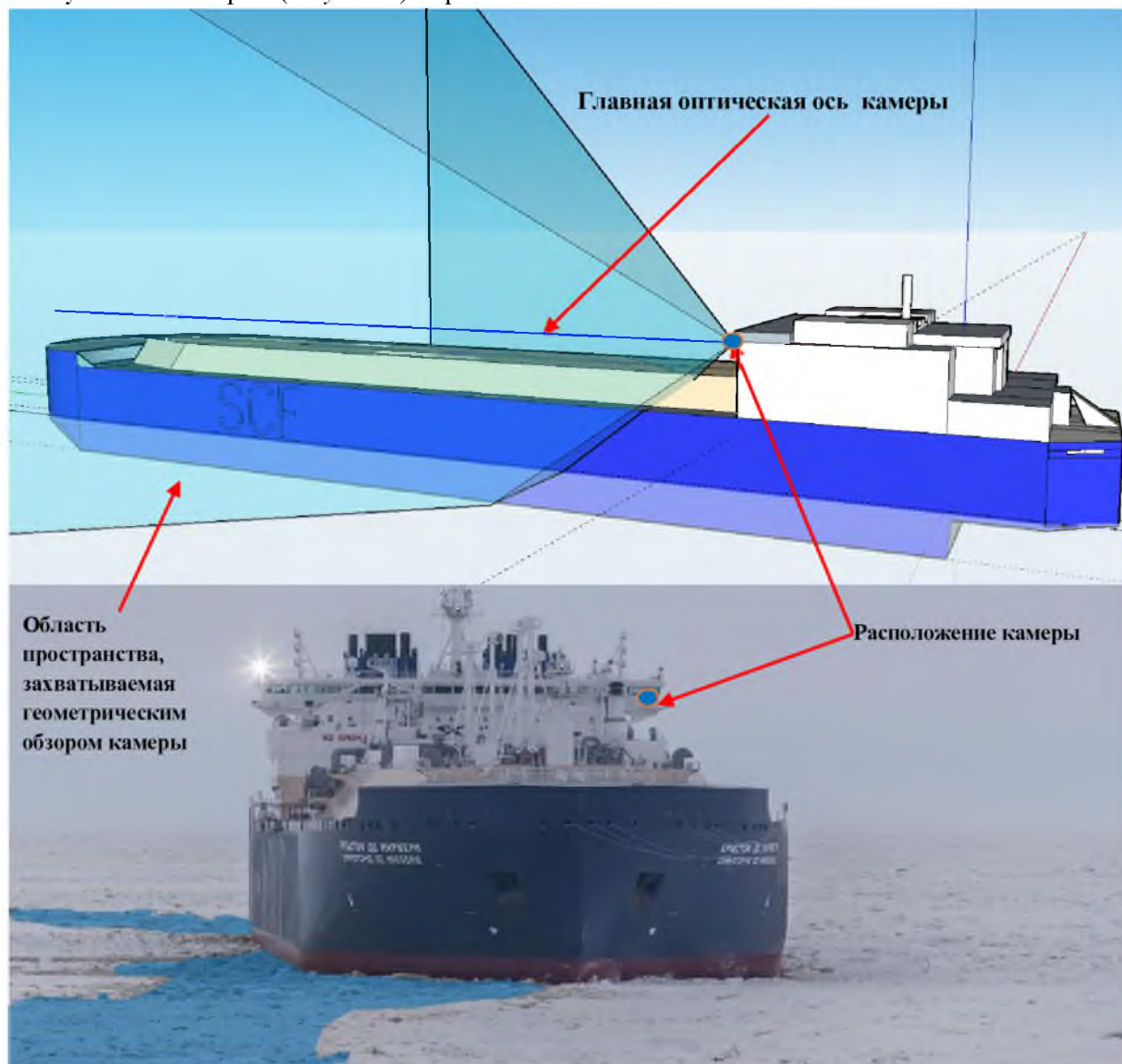


Рисунок 1 – Расположение цифровой камеры для записи экспериментальных данных

Выбор архитектуры сети. В рамках решения поставленной задачи обратимся к такой среде для глубинного обучения, как «Caffe» (ConvolutionArchitectureForFeatureExtraction-свёрточная архитектура для извлечения признаков).

«Caffe» поддерживает различные типы машинного обучения, нацеленные в первую очередь на решение задач классификации и сегментации изображений, обеспечивает свёрточные нейронные сети (Рисунок 2). «Caffe» позволяет использовать готовые промышленные конфигурации нейронных сетей, прошедшие апробацию.

В работе используется такая готовая переносимая архитектура сверточной нейронной сети,

как «MobileNet» [7]. «MobileNet» – эффективная архитектура, которая уменьшает количество используемой памяти для вычислений и при этом сохраняет высокую точность предсказаний, используя в своей структуре способы оптимизации скорости работы алгоритма.

Особенностью данной сети является использование так называемой «Хсертпюн» архитектуры - компактной глубокой нейронной сети. Такая архитектура способствует значительному сокращению количества параметров в сверточной сети за счёт замены обычных свёрток на «depthwise separable convolutions» - разделимые по глубине свертки.

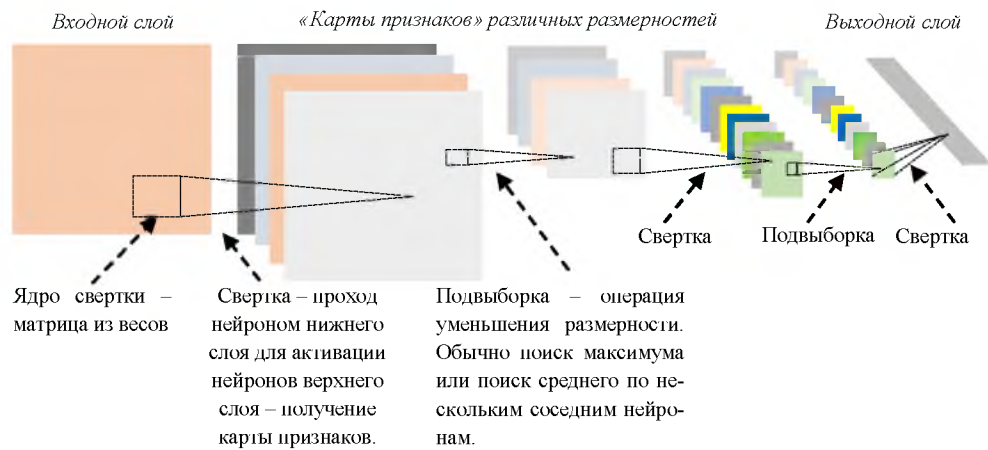


Рисунок 2 – Пример построения сверточной нейросети

Изначально «Xception» архитектура опирается на структурную архитектуру «Inception», идея которой заключается в том, что вместо выбора размера ядра свертки, используют одновременно несколько вариантов размера ядра и конкатенируют результаты. Данный подход увеличивает количество операций, необходимых для вычислений и активаций одного слоя. Чтобы упростить

вычислительную сложность, архитектура «Inception» предлагает перед каждым сверточным блоком выполнять свертку размером ядра 1x1. Это способствует снижению размерности сигнала, который подается на вход сверткам с большими размерами ядер (Рисунок 3).

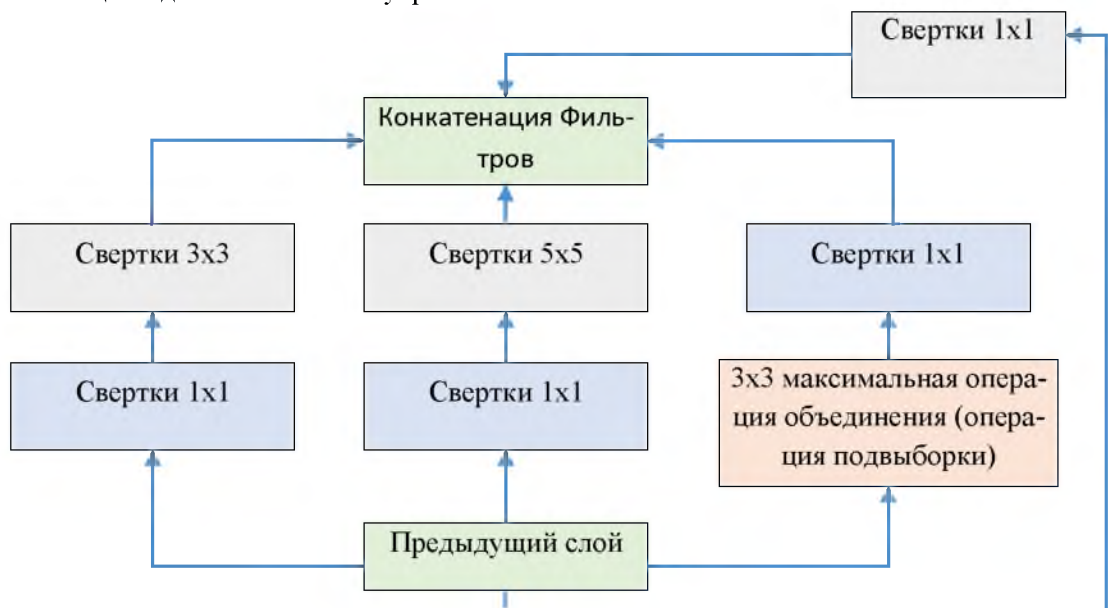


Рисунок 3 – Модуль «Inception»

В 2016 году была предложена «Xception» архитектура [4] (использование «экстремального» Inception модуля), получившего название «depthwise separable convolutions». Его основная идея заключается в следующем:

Пусть имеется стандартный сверточный слой с K_2 фильтрами размера 3x3. На вход в слой подается тензор размерности:

$$M * N * K_1, \quad (1)$$

где M, N это соответственно ширина и высота тензора, а K_1 это количество каналов. Тогда такой слой свернет одновременно все каналы исходного

сигнала K_2 различными свертками, а на выходе у данного слоя получится тензор размерности:

$$(M - 2) * (N - 2) * K_2, \quad (2)$$

Однако, если сделать вместо выше представленных расчетов два последовательных шага, то сеть станет компактнее:

- Произвести «поточечную» свертку - свернуть исходный тензор 1x1 сверткой, подобно Формуле 2.

- Произвести пространственную свертку «по глубине» - свернуть отдельно каждый канал сверткой 3x3 (из-за того, что сворачиваются не

все каналы вместе, общаеразмерность не изменится).

Например, свернем изображение с 8 каналами сверточным слоем 16 фильтрами. Суммарно такой сверточный слой получит $8 * 16 * 3 * 3 = 1152$ весов, так как имеем $8 * 16$ свёрток размерности 3×3 .

Если же в аналогичном примере применить «depthwise separable convolutions» - разделимые по глубине свертки, тогда:

Количество весов рассчитывается как $8 * 16 * 1 * 1 = 128$ у «поточечной» свертки. И $16 * 3 * 3 = 144$ веса у пространственной свертки «по глубине». В сумме получится 272 весов, что на порядок меньше количества весов обычного сверточного слоя (1152). Соответственно, требуется и меньшая вычислительная мощность.

В «MobileNet», считая поточечную и пространственную свертку за отдельный слой, насчитывается всего 28 слоев. Согласно официальной информации из оригинальной статьи разработчика [7], при обучении на датасете «ImageNet» точность применения модели «MobileNet» составляет порядка 70,6 %, а общее количество структурных параметров сети – 4,2 млн.

Обучение. Итак, первым этапом, для обучения нейронной сети, было подготовлено около 400 изображений (кадров), выделенных из различных нарезок видеоряда. На каждом из этих

изображений было выделено от 20 до 40 фрагментов разлома (лед - вода).

С помощью глубокой нейронной сети «Caffe», на модели «MobileNet», было проведено обучение с использованием имеющегося набора данных (датасета). После 12000 итераций (~50 часов), был достигнут приемлемый результат ошибки < 2%.

Далее обученная модель была загружена для обработки в программу, написанную в программной среде Python.

Обработка видеоряда происходит по кадрам, то есть алгоритм на каждом последовательном кадре находит фрагменты разлома (граница сплошного льда и ледового «трека») и записывает координаты в заранее созданный массив.

После наполнения массива 50 координатами фрагментов, происходит передача этого массива в функцию, которая выполняет необходимую математическую обработку и сортировку массива. Ясно, что детектированные координаты не могут находиться на одной линии из-за фактической неровности краев ледового «трека» (Рисунок 4). Необходимо вычислить теоретическую линию, соответствующую общему направлению ледового разлома. Поэтому, на основании массива данных с детектированными координатами разлома, выстроим предполагаемую линию разлома с помощью линейной регрессии [6].

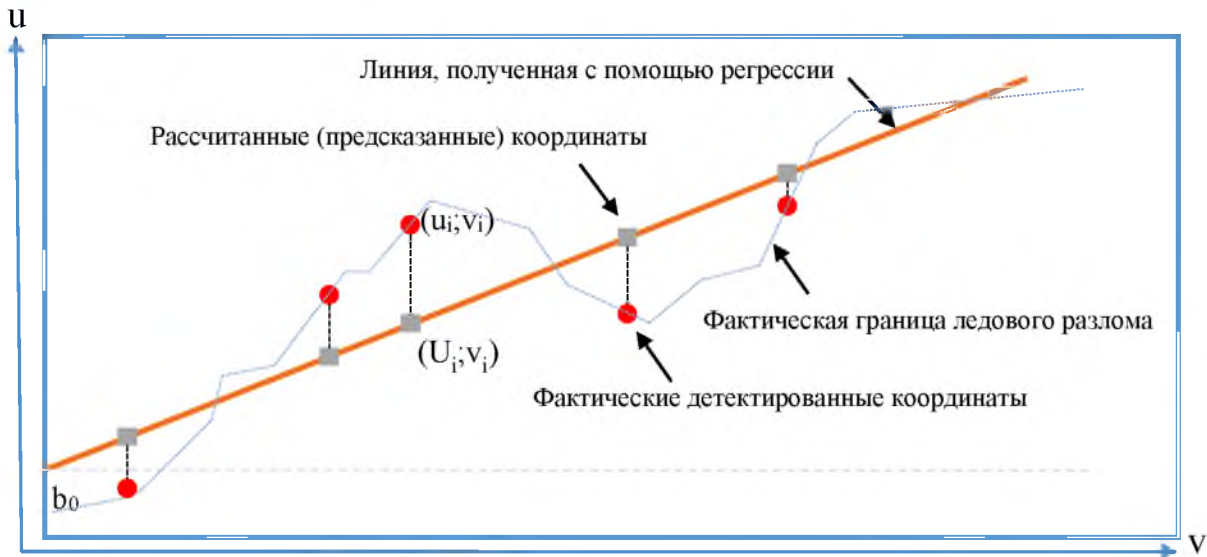


Рисунок 4 – Поиск теоретической линии ледового «трека»

Пусть пара (u_i, v_i) – пиксельные значения координат. Тогда U_i - предсказываемое значение исходной переменной при одинаковом v_i . Уравнение прямой зададим элементарно как функцию:

$$U_i = av_i + b_0 \quad (3)$$

Ошибку функции следует рассчитать как сумму отклонений фактического значения от расчётного:

$$\sum_{i=1}^N = (u_i - U_i)^2 \quad (4)$$

где $N = 50$ по условию алгоритма. Полученное выражение можно интерпретировать как функцию ошибок. Тогда можно найти значения, при которых частные производные равны нулю:

$$\frac{dE}{da} = 0; \frac{dE}{db} = 0 \quad (5)$$

По переменной a частная производная равна:

$$\frac{dE}{da} = \sum_{i=1}^N 2(u_i - U_i) \quad (6)$$

Тогда можем переписать выражение относительно переменной a , используя *Формулу 4*:

$$\frac{dE}{da} = \sum_{i=1}^N 2(u_i - (av_i + b))(-v_i) = 0 \quad (7)$$

Приравняв производную к 0, получим следующее выраженное равенство:

$$a \sum_{i=1}^N v_i^2 + b \sum_{i=1}^N v_i = \sum_{i=1}^N u_i v_i \quad (8)$$

Теперь относительно переменной b частная производная равна:

$$\frac{dE}{db} = \sum_{i=1}^N 2(u_i - (av_i + b))(-1) = 0 \quad (9)$$

Получим после упрощения:

$$a \sum_{i=1}^N v_i + b \sum_{i=1}^N 1 = \sum_{i=1}^N u_i = \sum_{i=1}^N v_i + b N \quad (10)$$

Имеем 2 уравнения (*Формула 8 и 10*) с 2 неизвестными a и b . После выражения и сокращения получим:

$$a = \frac{N \sum_{i=1}^N u_i v_i - \sum_{i=1}^N v_i \sum_{i=1}^N u_i}{N \sum_{i=1}^N v_i^2 - (\sum_{i=1}^N v_i)^2} \quad (11)$$

$$b = \frac{\sum_{i=1}^N u_i \sum_{i=1}^N v_i^2 - \sum_{i=1}^N v_i \sum_{i=1}^N u_i v_i}{N \sum_{i=1}^N v_i^2 - (\sum_{i=1}^N v_i)^2} \quad (12)$$

Упростим полученные выражения – найдем средние значения. Средняя величина переменной определяется как сумма всех таких переменных, делённая на их количество:

$$\bar{v}_i = \frac{1}{N} \sum_{i=1}^N v_i \quad (13)$$

Аналогично справедливо равенство:

$$\bar{u}_i \bar{v}_i = \frac{1}{N} \sum_{i=1}^N v_i u_i \quad (14)$$

Следовательно, можно переписать выражения относительно a и b :

$$a = \frac{\bar{u}_i \bar{v}_i - \bar{v}_i \bar{u}_i}{v_i^2 - \bar{v}_i^2} \quad (15)$$

$$b = \frac{\bar{u}_i v_i^2 - \bar{v}_i \bar{u}_i \bar{v}_i}{v_i^2 - \bar{v}_i^2} \quad (16)$$

После передачи массива для вычислений, аналогичных вышепредставленным, стек обнуляется, и наполнение массива начинается на основании новых данных.

Координаты линии также записываются в отдельный массив, для дальнейшего нахождения угла наклона линии (относительно системы координат кадра).

Как было упомянуто ранее, за истинный курс судна принимаем вертикальную линию левого борта судна, расположенную по центру видео, так как цифровая камера направлена строго по курсу.

Таким образом, в результате работы алгоритма находим направление линии ледового «трека» в «экранной» проекции. Ясно, что цифровая камера проецирует изображение реальной сцены с искажением.

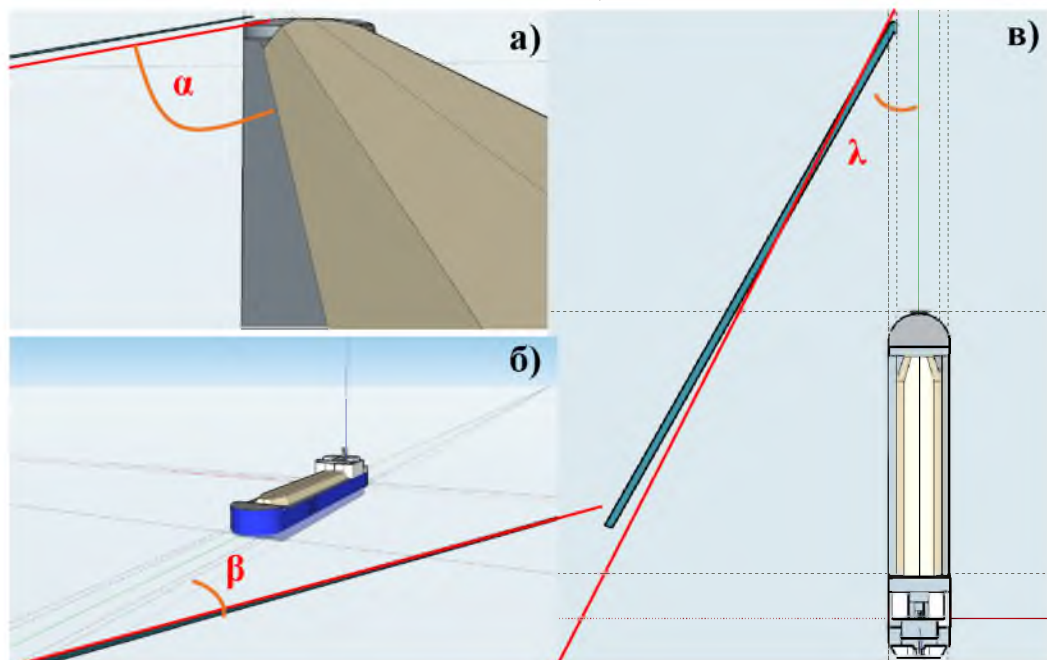


Рисунок 5 – Визуальное изменение значения отклонения курса судна от направления ледового «трека» в зависимости от изображения сцены: а) вид с левого крыла ходового мостика, б) произвольный вид со стороны, в) вид сверху – перспективная проекция

Поиск соотношения между углами на изображении сцены и их реальными значениями – сложная задача, которая требует обращения к такому разделу, как компьютерная графика. На *Рисунке 5* формализована проблема искажения углов. Так как размеры углов при перспективной проекции (вид сверху) будут фактически соответствовать углам, используемым в стандартных проекциях для навигации, то угол λ и будет искомым:

$$\lambda = \alpha * k \quad (17)$$

где α – угол, полученный путем поиска линии на изображении, k – сложный коэффициент перевода значений угла между проекциями.

В общем, при решении задачи поиска фактического значения коэффициента k для каждого случая расположения линии ледового «трека», применялись: калибровка камеры, использование информации о реальных размерах объектов на имеющихся кадрах (корпус судна), соотношения

перевода сцены кадра в перспективную проекцию. Также не обошлось без использование тригонометрических формул. Опустим данные выкладки, так как это не является предметом настоящей статьи. Главное – это полученное значение фактического угла отклонения курса судна от ледового «трека», выдаваемое в градусах.

На *Рисунке 6* представлен общий вид рабочей программы с обозначением детектированных точек ледового разлома, отрисованной линией и выведенным на экране фактическим значением угла отклонения от ледового «трека».

Стоит также заметить, что при работе программы, рассчитанный градус отклонения появляется не сразу. Это связано с особенностью работы алгоритма, которому необходимо для верного определения накопить полученные данные. Скорость накопления этих данных, зависит от количества распознанных фрагментов разлома.

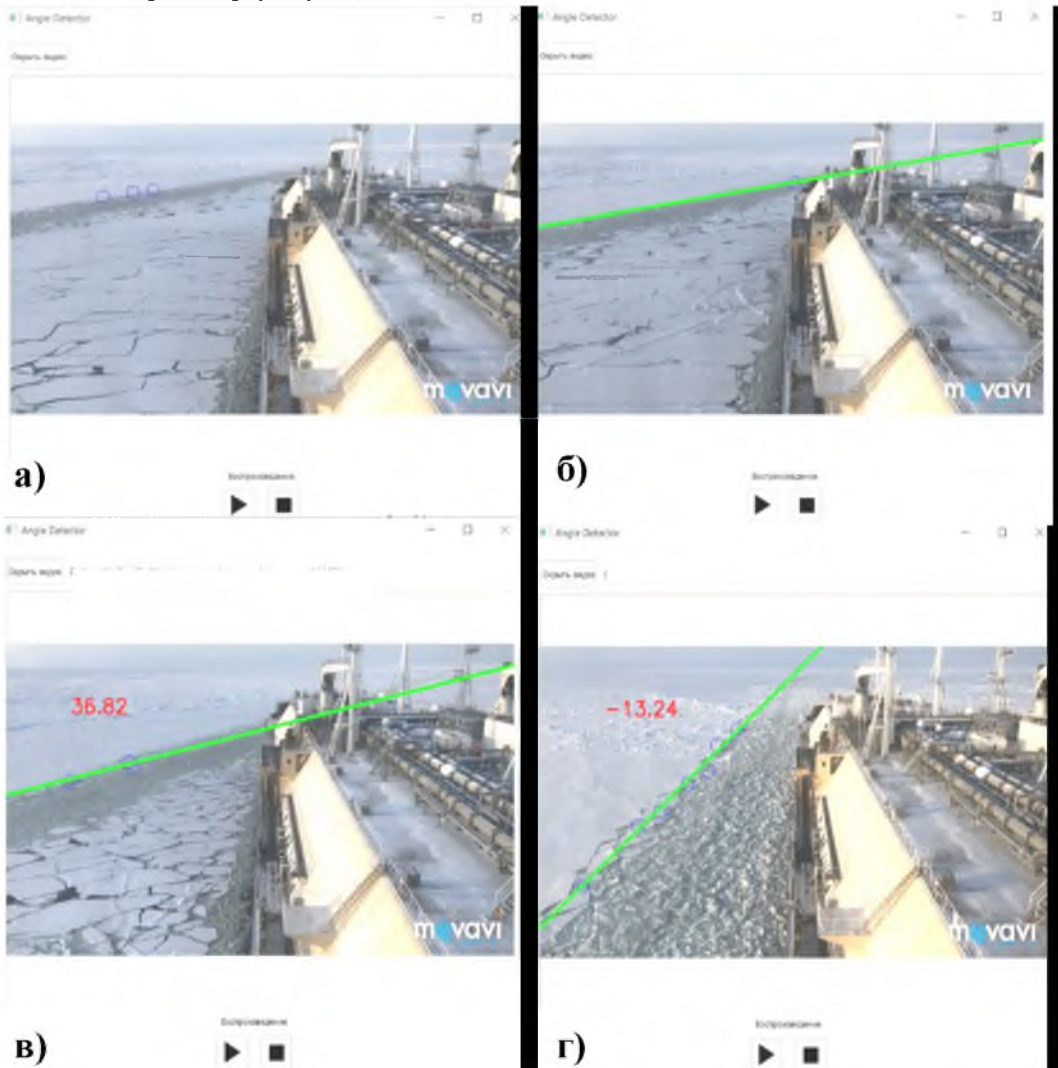


Рисунок 6 – Окно программы по автоматическому поиску направления ледового «трека». а) – начало работы программы, детектирование точек разлома (синие круги) обученной нейронной сетью, б) расчет и отрисовка теоретической линии ледового «трека» после накопления массива детектированных точек разлома, в) стабильная работа программы, вывод на экран значения угла отклонения ледового «трека» от истинного курса судна, г) другой фрагмент видеоряда (другое время съемки, освещенность сцены, тип ледового разлома) - стабильная работа программы

После тестирования работы программы на различных фрагментах видеоряда была отмечена стабильная работа алгоритма. При различных вариантах освещенности сцены и структуры ледяного разлома значение угла отклонения курса от ледового «трека» рассчитывается верно. Однако, из-за того, что камера установлена на крыле судна, четкое определение ледового разлома становится возможным только в ближней и средней перспективе поля зрения сцены (около 2/3 пространства от нижней рамки кадра). Несмотря на попытки, обучить нейросеть на нахождение фрагментов разлома в верхней 1/3 части кадра, распознавание в этой зоне происходит очень редко. Связано это с тем, что разрешение в этой части изображения очень маленькое, и настоящий алгоритм не позволяет каким-либо образом улучшить работу нейросети в данной области. Программное решение данной проблемы весьма сложное и может способствовать значительному усложнению алгоритма, увеличению вычислительных мощностей. Тем не менее, технически решить данную проблему можно при использовании цифровой камеры, установленной на баке судна. Информация с такой камеры позволит увеличить дальность обнаружения ледовых разломов.

Заключение. Сверточная нейронная сеть, предназначенная для определения границы ледового «трека», прошла процесс обучения. Испытания на различных фрагментах видео-ряда показали стабильную работу и достаточную точность вычислений. Созданная программа применима к практическому использованию на судах при ледовой навигации для автоматизации процесса поиска ледовых «треков», в том числе и для использования на СПГ Газовозах типа Arc7. В отличие от визуального поиска ледовых разломов судоводителем, программное средство, рассматриваемое в работе и использующее нейронную сеть, осуществляет расчет направления ледового «трека» полностью автоматизированным образом, что можно рассматривать в качестве нового технического средства помощи судоводителю в рамках эргатических систем управления. С другой стороны, развитие и интеграция данного программного средства с другими судовыми системами - это реальная перспектива внедрению мало и без-экипажного судовождения, а также полностью автоматических навигационных судовых систем.

Литература

- 1) Гиниятова Д. Нейронные сети на python. – Казань: Казанский Федеральный Университет,

2017. – 7 с. Режим доступа: https://kpfu.ru/portal/docs/F_1458204831/Nejronnye.seti.na.Python.pdf
- 2) Гонсалес Р., Вудс Р. Цифровая обработка изображений // Техносфера. – 2005. – 1007 с.
- 3) Ильин И.В., Гудков К.В. Анализ программных средств для глубинного обучения искусственных нейронных сетей. – Пенза: Пензенский государственный технологический университет, 2018. – 7с.
- 4) Созыкин А. В. Обзор методов обучения глубоких нейронных сетей // Вестник Южно-Уральского Гос. Университета. – 2017. – Том 6. – № 3 – 31 с.
- 5) Тутьчинский В.В., Тутьчинский В.И. Рассмотрение возможности применения элементов машинного зрения для автоматизации ледовой проводки на примере СПГ газозовов типа Arc7. // Эксплуатация морского транспорта. – 2019. – №3. – С. 67-76.
- 6) Шапков В.Б. Прикладной Регрессионный анализ. – Оренбург: Оренбургский государственный университет, 2003. – 363 с.
- 7) Howard A.G., ZhuBoChen M., Kalenichenko D., Wang W., Weyand T., Andreetto M., Hartwig A. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Google Inc., 2017. – 9 p.

References

- 1) Giniyatova D. Nejrornyie seti na python. Kazanskij Federal'nyj Universitet, 2017. – 7 s. Rezhim-dostupa: https://kpfu.ru/portal/docs/F_1458204831/Nejronnye.seti.na.Python.pdf
- 2) Gonsales R., Vuds R. Cifrovaya obrabotka izobrazhenij. Tekhnosfera. 2005. – 1007 s.
- 3) Il'in I. V., Gudkov K. V.. Analiz programmyh sredstv dlya glubinnogo obucheniya iskusstvennyh nejronnyh setej. Penzenskij gosudarstvennyj tekhnologicheskij universitet, 2018. – 7s.
- 4) Sozykin A. V. Obzor metodov obucheniya glubokih nejronnyh setej. Vestnik YUzhno-Ural'skogo Gos. Universiteta. Tom 6, № 3 2017. – 31 s.
- 5) Tul'chinskij V.V., Tul'chinskij V.I. Rassmotrenie vozmozhnosti primeneniya elementov mashinogo zreniya dlya avtomatizacii ledovoj provodki na primere SPG gazovozov tipa Arc7. GMU im. adm. F.F. Ushakova, nauchnyj zhurnal «Ekspluatatsiya morskogo transporta» vypusk №3, 2019. s. 67-76.
- 6) SHashkov V.B. Prikladnoj Regressionnyj analiz. Gosudarstvennoe obrazovatel'noe uchrezhdenie vysshego professional'nogo obrazovaniya «Orenburgskij gosudarstvennyj universitet» 2003. – 363 s.
- 7) Howard A.G., ZhuBoChen M., Kalenichenko D., Wang W., Weyand T., Andreetto M., Hartwig A. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Google Inc., 2017. – 9 p.