

ВЛИЯНИЕ АДДИТИВНЫХ ПОМЕХ НА КАЧЕСТВО СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ СНИМКОВ ДЕРЕВЬЕВ ПИХТЫ, ВЫПОЛНЯЕМОЙ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ МО-U-NET

Булько С.В.¹, Малкин А.Ю.²

¹НИ ТПУ, ИШИТР, студент, гр. 8K12, e-mail: svb44@tpu.ru

²НИ ТПУ, ИШИТР, аспирант, e-mail: aym27@tpu.ru

Аннотация

Исследуется влияние аддитивного гауссового шума на качество семантической сегментации снимков деревьев пихты сибирской, проводимой с помощью модели сверточной нейронной сети Мо-U-Net. Показано, что наличие аддитивных шумов на обучающих фрагментах датасета ухудшает качество семантической сегментации тестового снимка.

Ключевые слова: беспилотный летательный аппарат, семантическая сегментация снимков деревьев пихты сибирской, аддитивный гауссов шум, помехоустойчивость сверточной нейронной сети

Введение

В настоящее время специалисты лесной отрасли вынуждены проводить оперативный мониторинг пораженных уссурийским полиграфом деревьев пихты сибирской на довольно значительной территории России. Очаги такого вредителя пихтовых лесов сегодня выявлены в Томской области, Удмуртии, Прибайкалье и т. д. [1, 2]. Для оперативного мониторинга применяются (БПЛА), оснащенные устройствами цифровой фото- и (или) видеосъемки. Для автоматической семантической сегментации (попиксельной классификации) на таких снимках деревьев пихты по классам здоровья (степени поражения уссурийским полиграфом) активно применяются сверточные нейронные сети (СНС) [3]. В результате работы аппаратуры БПЛА на снимках могут возникать аддитивные помехи (часто говорят - шумы), что влияет на качество семантической сегментации (на точность попиксельной классификации) деревьев на снимках моделями СНС. Для снижения влияния аддитивного шума на качество сегментации в моделях СНС применяются функции потерь, являющиеся робастными: они обладают определенной устойчивостью к шумам. Примером таких функций является робастная функция потерь Уэлша [4], которая может быть применена при обучении выбранной нами для исследования модели СНС Мо-U-Net [5].

В данной работе изучается устойчивость модели СНС Мо-U-Net, обученной с робастной функцией потерь Уэлша, к аддитивному гауссовому шуму на снимках с БПЛА пораженных уссурийским полиграфом деревьев пихты сибирской.

Подготовка датасета

Для обучения, валидации и исследования выбранной и описанной ниже модели СНС необходимо иметь соответствующие наборы данных.

В качестве исходного набора данных используется датасет, подготовленный в рамках исследования [3]. Исходные снимки, полученные в [3] с помощью фотокамеры на БПЛА, были размечены авторами этой статьи с учетом следующих классов здоровья деревьев пихты сибирской: «Живые», «Отмирающие», «Свежий сухостой» и «Старый сухостой». Кроме того, был объявлен класс «Фон», куда входят деревья, не относящиеся к деревьям породы пихты сибирской, или прочие объекты на снимке. По размеченным снимкам датасета подготовлены обучающая выборка, которая состоит из 2004 фрагментов, а также валидационная выборка – 672 фрагмента и тестовая выборка – 91 фрагмент. Все фрагменты каждого RGB - канала имеют размер 256x256 пикселей.

На основе исходного датасета нами создается новый датасет, содержащий фрагменты с искусственным образом вводимыми аддитивными помехами. В новом датасете зашумляется 25% площади обучающих фрагментов, то есть 25% пикселей от всего количества пикселей каждого фрагмента. Пиксели фрагмента, подлежащие зашумлению, выбираются случайным образом по дискретному равномерному закону распределения. Измененные значения амплитуд яркости каждого из RGB-каналов $\omega(x,y)$ зашумленного пикселя вычисляются по формуле:

$$\omega(x,y) = s(x,y) + n(x,y), \quad (1)$$

где x и y – координаты пикселя, на который будет накладываться шум; $s(x,y)$ – начальные (исходные) значения амплитуды яркости пикселей каждого из RGB-каналов; $n(x,y)$ – случайная амплитуда накладываемого на полезный сигнал шума (изменения значений амплитуды яркости пикселей RGB-каналов).

Пусть амплитуда аддитивной помехи представлена случайной величиной, полученной из распределения Гаусса (нормальное распределение), задаваемого функцией плотности вероятности:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(z-\mu)^2}{2\sigma^2}}, \quad (2)$$

где z – значение амплитуды помехи, вычисляемое как вероятность включения этого значения в диапазон $[\mu-\sigma; \mu+\sigma]$; μ – математическое ожидание случайной величины, σ – ее среднеквадратичное отклонение.

На рисунке 1 представлены примеры зашумления фрагмента снимка деревьев пихты аддитивными помехами с различными параметрами. Визуальный анализ этих примеров позволяет считать, что наличие таких помех с определенными значениями параметров на фрагменте оказывает некоторое влияние на его качество.

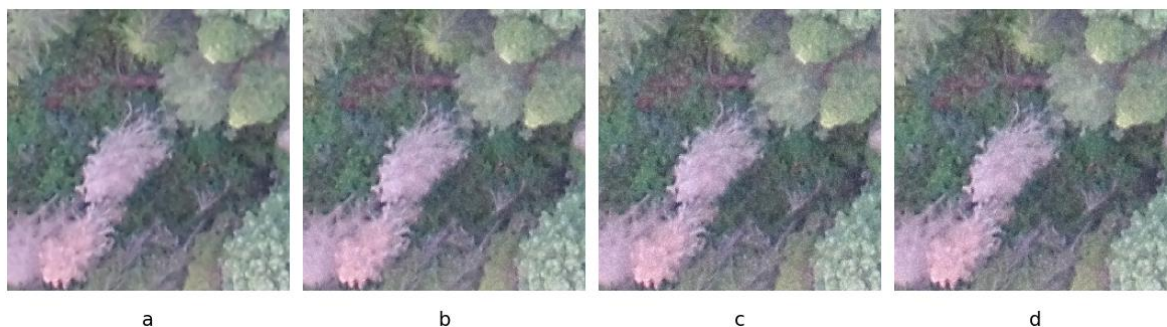


Рис. 1. Зашумление фрагмента снимка деревьев пихты: а – исходный фрагмент; наложение аддитивного шума, математическое ожидание μ и среднеквадратичное отклонение которого, соответственно, составляют: b – 15 и 0,5; c – 20 и 0,5; d – 20 и 2.0

Задача классификации снимков с аддитивными помехами

Пусть для решения задачи семантической сегментации снимков с БПЛА деревьев пихты сибирской, пораженных уссурийским полиграфом, применяется модель СНС Mo-U-Net [5]. Эта модель является модификацией классической архитектуры U-Net [6]. Особенностью модели Mo-U-Net является сочетание пространственной информации и контекстных признаков, что показало эффективность ее в решении задач семантической сегментации (попиксельной классификации) снимков деревьев пихты.

При обучении модели будет применяться робастная функция потерь Уэлша, определяемая в [4] как:

$$L = 1 - \exp\left(-\frac{1}{2}\left(\frac{k}{\beta}\right)^2\right), \quad (3)$$

где β – коэффициент функции потерь; k – величина ошибки предсказания модели. Производная данной функции имеет вид:

$$L' = \frac{1}{\beta^2} k \exp \left(-\frac{1}{2} \left(\frac{k}{\beta} \right)^2 \right). \quad (4)$$

Выражение (4) является непрерывным во всем множестве вещественных чисел, что позволяет использовать (3) в качестве функции потерь в алгоритме обратного распространения ошибки при обучении модели СНС. В проводимом эксперименте будут браться значения коэффициента β в диапазоне [0,001; 9,000].

Проводится серия экспериментов, на первом этапе которых создается новый датасет. Для этого искусственно зашумляются обучающие фрагменты исходного датасета с фиксированной площадью зашумления, равной 25%, при различных значениях параметров нормальной (гауссовской) плотности распределения вероятности. Так параметр μ , принимает значения 10, 15 и 20 при фиксированном среднеквадратичном отклонении распределения вероятности σ , равном 2,0. Затем готовятся фрагменты нового датасета для исследования влияния σ , принимающего значения 0,5, 1,0, 1,5, при фиксированном математическом ожидании μ , равном 20. Результаты исследования эффективности модели Mo-U-Net в этих экспериментах будут также сравниваться с результатами обучения, валидации и тестирования этой модели на исходном датасете из [3], то есть на датасете без аддитивных помех.

При очередных значениях параметров аддитивных помех в течение 50 итераций с применением байесовского алгоритма оптимизации определяется квази-оптимальное значение коэффициента робастной функции потерь Уэлша, при котором обученная и валидированная модель будет демонстрировать максимальное значение метрики mIoU на валидационной выборке. Такая модель затем будет исследоваться на тестовой выборке датасета.

Для оценки точности классификации снимков в качестве критерия точности модели использовались следующие метрики. Во – первых, это IoU_c (Intersection over Union for Class) – мера оценки степени перекрытия между предсказанными и истинными областями для одного класса c деревьев:

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c}, \quad (5)$$

где True Positive c (TP_c) – количество пикселей, которые правильно предсказаны как принадлежащие классу c ; False Positive c (FP_c) – количество пикселей, которые предсказаны как принадлежащие классу c , но на самом деле принадлежат другому классу; False Negative c (FN_c) – количество пикселей, которые должны принадлежать данному классу, но были классифицированы как принадлежащие другому классу.

Еще одной метрикой является Mean Intersection over Union – мера оценки степени перекрытия между предсказанными и истинными областям, усредненная по всем классам:

$$mIoU = \frac{1}{N} \sum_{c=1}^N IoU_c, \quad (6)$$

где N – количество классов деревьев пихты в датасете, равное пяти. Будем считать, что обученная и валидированная модель Mo-U-Net является эффективной в решении задачи мультиклассификации деревьев пихты на снимках, если в результате сегментации тестового снимка моделью значения метрик IoU_c и $mIoU$ превышают 0,5. Кроме того, будем считать метрику $mIoU$ мерой устойчивости исследуемой модели Mo-U-Net к помехам.

На языке Python с применением фреймворка PyTorch программно реализована модель СНС Mo-U-Net, а также робастная функция потерь Уэлша. Проведены обучение и валидация модели на новом датасете, получены следующие результаты.

Результаты экспериментов

Алгоритм проведения каждого из экспериментов следующий. Покажем его на конкретном примере. В результате обучения и валидации модели Mo-U-Net на фрагментах,

содержащих аддитивный шум, математическое ожидание которого составляет 10, а среднеквадратичное отклонение – 2,0, максимальное значение метрики mIoU было достигнуто при значении коэффициента β РФП Уэлша, равном 1,4453. В этом случае значение метрики mIoU равно 0,6751. На рисунке 2 представлен ход эксперимента по подбору квази-оптимального значения коэффициента β функции потерь в рассматриваемом случае. Из рисунка следует, что квази-оптимальное значение коэффициента β РФП Уэлша было получено на 22 итерации проводимого эксперимента при достижении максимального значения метрики mIoU, равного 0,6751.

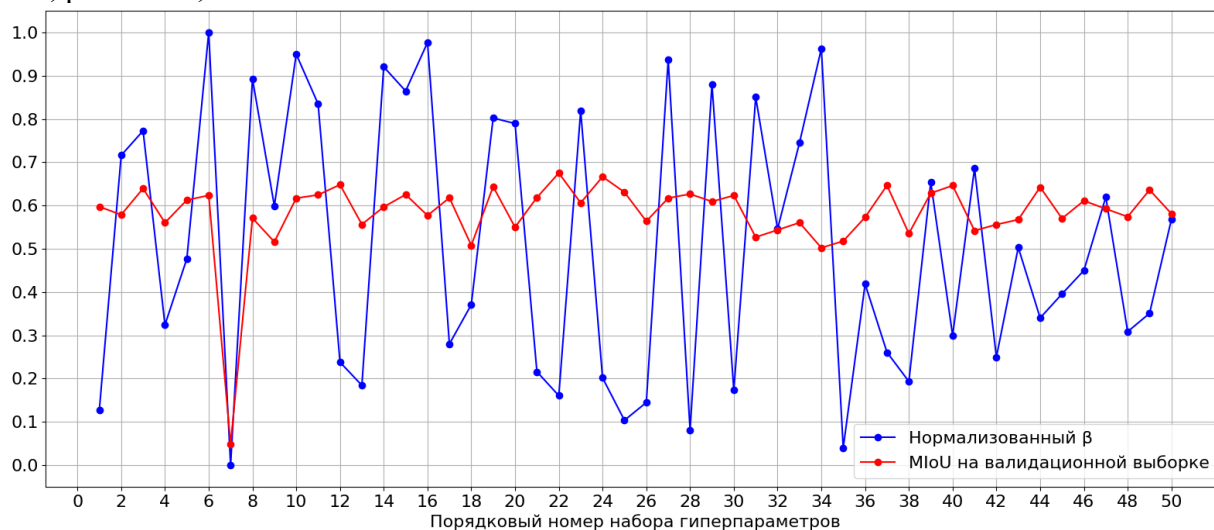


Рис. 2. Валидация модели, обученной на зашумленных аддитивными помехами данных с математическим ожиданием, равным 10, на фрагментах

Получены аналогичные результаты при анализе подобных рисунков для других рассматриваемых сценариев зашумления обучающих фрагментов. Так, при математическом ожидании аддитивных помех на обучающих фрагментах, равном 15, квази-оптимальное значение коэффициента β составляет 7,1352. В таком случае метрика mIoU на валидационной выборке примет значение 0,7057. При величине математического ожидания аддитивных помех, равном 20, квази-оптимальным значением коэффициента функции потерь является 3,7540. Значение метрики mIoU на валидационной выборке при этом составляет 0,7048.

При обучении модели на фрагментах, содержащих аддитивные шумы с различным среднеквадратичным отклонением при постоянном математическом ожидании, равном 20, были получены квази-оптимальные значения коэффициента β функции потерь: 1,9193, 3,9096, 5,1552 при σ , равном, соответственно, 1,5, 1,0, 0,5. В таких случаях моделью Mo-U-Net получены следующие значения метрики mIoU на валидационной выборке: 0,6914, 0,6968, 0,7404. При обучении модели на фрагментах без аддитивных помех квази-оптимальное значение коэффициента β функции потерь Уэлша равно 5,8181. Метрика mIoU на валидационной выборке в данном случае составляет 0,7156.

Обученная и валидированная модель Mo-U-Net была протестирована на тестовой выборке. Все полученные результаты тестирования в виде значений метрики mIoU представлены в таблице 1. Из нее следует, что наличие аддитивных помех на обучающих фрагментах уменьшает точность классификации валидационных и тестовых фрагментов. Однако, максимальное уменьшение значения метрики mIoU на тестовой выборке составляет 0,0668 относительно случая отсутствия аддитивных помех и является небольшим. Стоит обратить внимание, что довольно значительное изменение параметров μ и σ нормальной (гауссовской) плотности распределения вероятности помех не оказывает значимого влияния на точность классификации деревьев пихты на снимках, выполняемой с помощью модели SNC Mo-U-Net.

Таблица 1. Квазиоптимальные значения коэффициента функции потерь модели и значения метрики mIoU

Параметры аддитивного шума обучающих фрагментов		Коэф-т функции потерь	mIoU на валидационной выборке	mIoU на тестовой выборке
μ	σ			
отсутствует аддитивный шум		5,8181	0,7156	0,6864
10	2,0	1,4453	0,6751	0,6272
15	2,0	7,1352	0,7057	0,6196
20	2,0	3,7540	0,7048	0,6235
20	1,5	1,9193	0,6914	0,6334
20	1,0	3,9096	0,6968	0,6309
20	0,5	5,1552	0,7404	0,6274

Заключение

Проведено исследование влияния аддитивных помех на фрагментах снимков деревьев пихты сибирской при обучении модели СНС Mo-U-Net в случае различных значений параметров аддитивных помех. Было установлено, что наличие аддитивных помех на обучающих фрагментах датасета оказывает влияние на качество сегментации тестового снимка в сторону его ухудшения. Тем не менее, в случае изменения как математического ожидания, так и среднеквадратичного отклонения нормальной (гауссовской) плотности распределения вероятности помех не приносит значительных изменений в точность классификации деревьев пихты на зашумленных аддитивными помехами снимках, получаемой с помощью модели Mo-U-Net. Это говорит о том, что РФП Уэлша способствует устойчивости исследуемой модели Mo-U-Net к аддитивным помехам для довольно больших диапазонов изменения параметров μ и σ нормального распределения случайных амплитуд этих помех.

Список использованных источников

1. Бисирова Э. М., Кривец С. А., Черногоров П. Н. Распространение и популяционные характеристики уссурийского полиграфа *Polygraphus proximus* на севере Томской области // Известия Санкт-Петербургской лесотехнической академии. – 2024. – №. 244. – С. 7-25.
2. Пашенова Н. В. и др. Чувствительность флоры хвойных Сибири к *Grosmannia aoshimae* (Ohtaka, Masuya & Yamaoka) Masuya & Yamaoka–симбионту уссурийского полиграфа // Дендробионтные беспозвоночные животные и грибы и их роль в лесных экосистемах (XI Чтения памяти О.А. Катаева). – 2020. – С. 247-248.
3. Керчев И. А. и др. Семантическая сегментация поврежденных деревьев пихты на снимках с беспилотных летательных аппаратов // Современные проблемы дистанционного зондирования Земли из космоса. – 2021. – Т. 18. – №. 1. – С. 116-126.
4. Сивак М. А. Робастное обучение нейронных сетей с простой архитектурой для решения задач классификации: диссертация на соискание ученой степени кандидата технических наук: 05.13.17. – Новосибирский государственный технический университет, Новосибирск, 2022. – 117 с.
5. Марков Н.Г., Мачука К. Модели и методы глубокого обучения для решения задач дистанционного мониторинга лесных ресурсов // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2024. – Т. 335. – № 6. – С. 55–74. DOI: 10.18799/24131830/2024/6/4600
6. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. – Springer international publishing, 2015. – С. 234-241.