

doi: 10.17586/2226-1494-2021-21-4-535-544

УДК 004.85

## Байесовские функции потерь для моделирования гомоскедастичной алеаторной неопределенности в задаче детекции пылицы на изображениях

Наталья Евгеньевна Ханжина

Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

Nehanzhina@gmail.com✉, <https://orcid.org/0000-0003-0795-642X>

### Аннотация

**Предмет исследования.** Представлены результаты моделирования гомоскедастичной алеаторной неопределенности для детекции пылевых зерен на изображениях. Приведено описание новых разработанных функций потерь моделирования неопределенности, основанных на фокусной и сглаженной  $L_1$  функциях потерь. Актуальность и отличие от существующих функций состоит в том, что базовые фокусная и сглаженная  $L_1$  функции, являясь передовыми для решения задачи детекции объектов на изображениях, не позволяют оценить алеаторную неопределенность. Разработанные функции позволяют дать такую оценку и представить более точное решение задачи детекции пылицы. **Метод.** Предложены функции потерь для обучения нейронной сети RetinaNet, моделирующие гомоскедастичную алеаторную неопределенность, для детекции объектов на изображении. Функции получены с помощью байесовского вывода и позволяют использовать его в существующих нейросетевых детекторах, основанных на архитектуре RetinaNet. Преимущества функций потерь продемонстрированы на задаче детекции изображений пылицы. **Основные результаты.** С помощью новых функций потерь удалось повысить точность детекции пылицы, а именно локализации и классификации, на изображениях в среднем на 2,76 %. Показано, что моделирование гомоскедастичной алеаторной неопределенности в процессе обучения нейронной сети позволяет отделять шум, присущий данным, от сигнала, тем самым повышая точность решения задач. **Практическая значимость.** Предложенные функции потерь позволяют значительно повысить точность детекции пылицы на изображениях, что имеет решающее значение для распознавания пылицы в целом. Результаты работы дадут возможность автоматизировать процесс определения пылицы аллергенов в воздухе и сократить время информирования больных поллинозами для предупреждения симптомов аллергии. Разработанные функции могут быть применены для обучения нейронной сети для детекции на любых других наборах данных изображений.

### Ключевые слова

байесовское глубокое обучение, байесовский вывод, алеаторная неопределенность, распознавание пылицы, детекция объектов на изображении, оценка неопределенности, байесовское моделирование

### Благодарности

Исследование выполнено при финансовой поддержке Национального центра когнитивных разработок Университета ИТМО. Автор выражает благодарность А.С. Лапенку, Н.В. Минаевой, Л.В. Новоселовой, Г.А. Заморину, Т.А. Полевой, А.А. Фильченкову, Е.Б. Замятиной, И.В. Харисовой, Ю. Пинаевой, Е.А. Цымбалову за неоценимую помощь в исследовании.

**Ссылка для цитирования:** Ханжина Н.Е. Байесовские функции потерь для моделирования гомоскедастичной алеаторной неопределенности в задаче детекции пылицы на изображениях // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21, № 4. С. 535–544. doi: 10.17586/2226-1494-2021-21-4-535-544

## Bayesian losses for homoscedastic aleatoric uncertainty modeling in pollen image detection

Natalia E. Khanzhina

ITMO University, Saint Petersburg, 197101, Russian Federation

Nehanzhina@gmail.com✉, <https://orcid.org/0000-0003-0795-642X>

### Abstract

The paper investigates the homoscedastic aleatoric uncertainty modeling for the detection of pollen in images. The new uncertainty modeling loss functions are presented, which are based on the focal and smooth  $L_1$  losses. The focal and smooth  $L_1$  losses proved their efficiency for the problem of image detection, however, they do not allow modeling the aleatoric uncertainty, while the proposed functions do, leading to more accurate solutions. The functions are based on Bayesian inference and allow for effortless use in existing neural network detectors based on the RetinaNet architecture. The advantages of the loss functions are described on the problem of pollen detection in images. The new loss functions increased the accuracy of pollen image detection, namely localization and classification, on average by 2.76 %, which is crucial for the pollen recognition in general. This helps to automate the process of determining allergenic pollen in the air and reduce the time to inform patients with pollinosis to prevent allergy symptoms. The obtained result shows that the modeling of homoscedastic aleatoric uncertainty for neural networks allows separating the noise from the data, increasing the accuracy of the proposed solutions. The developed functions can be applied to train neural network detectors on any other image datasets.

### Keywords

Bayesian deep learning, Bayesian inference, aleatoric uncertainty, pollen recognition, object detection, uncertainty quantification, Bayesian modeling

### Acknowledgements

This work is financially supported by National Center for Cognitive Research of ITMO University. The author would like to thank Alexey Lapenok, Natalia Minaeva, Larisa Novoselova, Georgiy Zamorin, Tatyana Polevaya, Andrey Filchenkov, Elena Zamyatina, Irina Kharisova, Yuliya Pinaeva and Evgeny Tsymbalov for their great help and useful comments.

**For citation:** Khanzhina N.E. Bayesian losses for homoscedastic aleatoric uncertainty modeling in pollen image detection. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 4, pp. 535–544 (in Russian). doi: 10.17586/2226-1494-2021-21-4-535-544

### Введение

Аллергия на пыльцу является глобальной проблемой для здоровья [1]. Более 20 % населения Европы имеет пыльцевую аллергию на различные виды растений [2]. В связи с высокой распространенностью респираторной аллергии данные о содержании в воздухе пыльцы аллергенных растений имеют важное медицинское и социально-экономическое значения [3]. Научный интерес представляет изучение закономерностей производства и распространения пыльцы аллергенных растений [4]. Данные о пыльце используются для решения многих актуальных вопросов аллергологии, аэриобиологии. Врачам важно определить пороговые уровни пыльцы, при которых развиваются клинические симптомы ринита, конъюнктивита, астмы [5, 6]. Исследователям необходимо точно документировать воздействие аллергена (пыльцы) при проведении клинических исследований, например, при аллергенспецифической иммунотерапии [7]. Существенные региональные различия содержания пыльцы обосновывают необходимость интеграции локальных данных и глобальной оценки проблемы [8, 9].

На этапе анализа пыльцы наиболее трудоемкая часть — идентификация видов растений для определения аллергенов или их распознавание. Этот шаг является ручным и занимает половину времени всего процесса анализа пыльцы, начиная со сбора проб из воздуха с пыльцевых ловушек, заканчивая тщательным распознаванием пыльцы под микроскопом [10, 11]. В зависимости от региона больные получают информацию с

7–10 дневной задержкой. Для ускорения процесса, с целью предоставления актуальной информации аллергикам, существует необходимость разработки инструментов для автоматизации этапа идентификации пыльцы.

Автоматизация распознавания изображений пыльцы может быть достигнута с помощью применения глубокого обучения, зарекомендовавшего себя в задачах компьютерного зрения как на этапе детекции пыльцевых зерен (которых на изображении может быть несколько), так и на этапе их классификации.

### Постановка задачи

Проблема автоматизации распознавания пыльцы была впервые выявлена в 1968 г. [12]. С тех пор исследователи по всему миру предпринимают попытки решить задачу с помощью различных методов. Многие выделяют узконаправленные признаки пыльцевых зерен, направленные на их природу, такие как форма, яркость, площадь, периметр, текстурные признаки, апертуры [13–17], и строят на их основе стандартные модели машинного обучения: метод опорных векторов, линейный дискриминантный анализ, случайный лес, нейросети,  $k$ -ближайших соседей и др. Автор данной работы также использовал стандартные методы машинного обучения для решения задачи распознавания пыльцы [18].

Несмотря на популярность и эффективность методов глубокого обучения, их применимость к задаче распознавания пыльцы впервые была исследована в работе [19]. Последующие результаты исследований по

этой теме целиком лежат в области глубокого обучения [20–22].

Основной недостаток большинства работ по данной задаче — игнорирование шага детекции пыли. Этот шаг является ключевым для автоматизации распознавания в целом, так как изображения с пылевой ловушки содержат, как правило, несколько пылевых зерен, а также объекты, не являющиеся пылью. Пример изображения пробы воздуха из пылевой ловушки, полученного с помощью оптического микроскопа приведен на рис. 1. Также трудность для построения моделей создает отсутствие открытых пылевых наборов данных, размеченных для задачи детекции.

Для решения указанной проблемы в данной работе использован частный набор данных, размеченный автором вручную. Так как разметка может иметь неточности, как это нередко бывает с наборами данных компьютерного зрения [23], для повышения точности решения задачи детекции пыли предложено использовать глубокое байесовское обучение.

Одним из способов учета ошибок разметки на основе байесовского вывода может служить оценка алеаторной неопределенности, которая отражает уровень шума в обучающей выборке, что позволит учитывать его на этапе предсказания. Алеаторная неопределенность делится на гомоскедастичную — однородную для всего распределения данных, и гетероскедастичную — различную для разных объектов данных. Несмотря на то, что оценка гетероскедастичной неопределенности полезнее для задач компьютерного зрения в целом [24], ее моделирование требует изменений в архитектуре нейронной сети. Также ее применение на практике требует разработки методов учета этой неопределенности при предсказании для конкретного объекта.

Как показывают исследования [25], моделирование гомоскедастичной алеаторной неопределенности может быть произведено на основе модификации лишь функций потерь, а не архитектуры в целом, что менее трудоемко. Также ее моделирование позволяет повысить точность решения задач [25] компьютерного зрения.

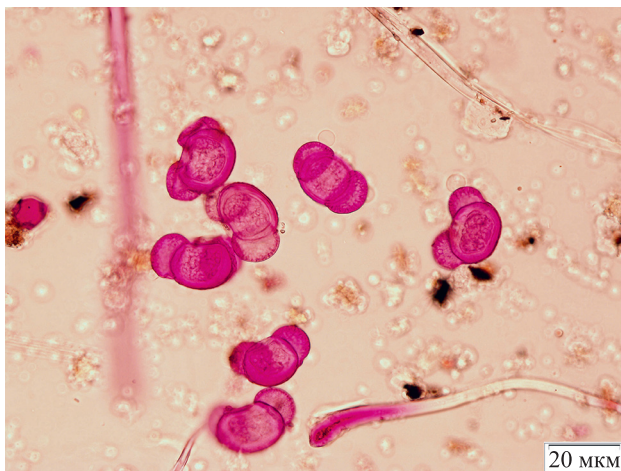


Рис. 1. Пример изображения пробы воздуха из пылевой ловушки, полученный с помощью оптического микроскопа  
Fig. 1. Example of an image of an air sample from a pollen trap taken with an optical microscope

Работа [25] рассматривает применение учета этого типа неопределенности для многозадачной архитектуры, решающей задачи семантической, инстанс-сегментации и предсказания глубины изображения.

В последнее время байесовское глубокое обучение широко используется для детекции объектов [26–32]. Однако все эти работы сосредоточены на другом типе неопределенности — эпистемической. Меньшее количество работ посвящено оценке алеаторной неопределенности [33, 34]. Но в существующих работах не изучается моделирование гомоскедастичной алеаторной неопределенности для задачи детекции, хотя это может помочь изолировать шум от данных и повысить надежность модели.

В настоящей работе предложены новые функции потерь, оптимизация которых эквивалентна моделированию гомоскедастичной алеаторной неопределенности для совместных задач локализации и классификации.

Представлены новые методы: байесовская фокусная функция потерь (Bayesian Focal Loss, BFL) — для оценки гомоскедастичной алеаторной неопределенности данных предсказания нейронной сети на основе байесовского вывода в задаче классификации и байесовская сглаженная  $L_1$  функция потерь (Bayesian Smooth  $L_1$  Loss) — в задаче локализации.

Предлагаемые функции потерь для моделирования гомоскедастичной алеаторной неопределенности использованы для обучения архитектуры RetinaNet [35]. Однако функции могут быть применены к любым нейросетевым детекторам, которые используют перекрестную энтропию (или фокусную функцию потерь для классификации), а также  $L_1$  (или сглаженную  $L_1$  функцию потерь для локализации), без изменения их архитектуры и процесса обучения.

Моделирование неопределенности поможет сделать существующие детекторы устойчивыми к шуму в разметке пылевых изображений, а также повысить точность детекции пыли.

### Обзор архитектуры RetinaNet для детекции изображений

RetinaNet [35] — одноэтапная якорная нейронная сеть для детекции объектов. Детекция включает в себя две задачи: локализацию — определение координат, ограничивающих объекты прямоугольников, и классификацию объектов интереса.

Локализация с точки зрения машинного обучения является задачей регрессии. Для регрессии ограничивающих прямоугольников RetinaNet использует сглаженную  $L_1$  функцию потерь. Это комбинация  $L_1$  и  $L_2$  функций, основанная на функции Хубера [36].

Его формула

$$\text{Smooth}L_1 = \begin{cases} \frac{\beta^2}{2} \varepsilon^2, & \varepsilon < \frac{1}{\beta^2}, \\ \varepsilon - \frac{1}{2\beta^2}, & \text{иначе} \end{cases},$$

где  $\frac{1}{\beta^2}$  — пороговое значение для перехода от  $L_1$  к  $L_2$  функции потерь;  $\varepsilon = \|y - f^W(x)\|$  — норма разницы меж-

ду истинной и предсказанной координатами,  $x$  — вход сети, ее выход для предсказанной координаты —  $f^W(x)$ ,  $y$  — истинная координата ограничивающего прямоугольника объекта. Основное отличие от функции потерь  $L_2$  заключается в том, что добавление интервала  $L_1$  помогает избежать чрезмерного штрафования нейросети за выбросы данных.

Для обучения задачи классификации в работе [35] предложена фокусная функция потерь, которая штрафует сеть лучше, чем перекрестная энтропия [35] на сложных негативных примерах (hard negative samples):

$$FL(p_i) = -(1 - p_i)^\gamma \log p_i,$$

где  $p_i = \begin{cases} p, & y = 1 \\ 1 - p, & \text{иначе} \end{cases}$ ,  $y$  — метка класса истинности объекта;  $\gamma$  — коэффициент модуляции;  $p = \text{Sigmoid}(f^W(x))$ .

Основное отличие фокальной функции потерь от перекрестной энтропии состоит в наличии коэффициента модуляции для решения проблемы дисбаланса классов. Данная проблема характерна для задачи детекции объектов, так как объект интереса ( $y = 1$ ) обычно занимает на изображении мало места. Таким образом, при больших ошибках классификации градиент фокусной функции потерь выше, чем у перекрестной энтропии, и наоборот. Это заставляет сеть лучше сосредоточиться на сложных негативных примерах.

Суммарная функция потерь RetinaNet имеет вид

$$L(f^W(x), y) = \alpha FL(f^W(x), y) + \text{Smooth}L_1(f^W(x), y),$$

где  $FL$  — фокусная функция потерь для классификации;  $\text{Smooth}L_1$  — сглаженная  $L_1$  функция потерь для локализации (регрессии);  $\alpha$  — коэффициент балансировки, который регулирует вклад  $FL$  в общую функцию потерь.

Несмотря на то, что функции потерь RetinaNet признаны передовыми для детекции, они не позволяют моделировать гомоскедастичную алеаторную неопределенность, что делает нейросеть чувствительной к шуму в данных. Для решения данной проблемы в настоящей работе предложены новые фокусная и сглаженная  $L_1$  функции потерь, которые способны моделировать гомоскедастичную алеаторную неопределенность. Для удобства нейронная сеть, использующая предложенные функции потерь, здесь и далее названа *байесовская RetinaNet*.

### Байесовская сглаженная $L_1$ функция потерь

На основании инструмента для моделирования гомоскедастичной алеаторной неопределенности для взвешивания многозадачных функций потерь [25], введем новую сглаженную  $L_1$  функцию потерь с гомоскедастичной неопределенностью, использующую оценку максимального правдоподобия.

Пусть  $f^W(x)$  — выход нейронной сети с весами  $W$  на входном изображении  $x$ ;  $\varepsilon$  — значение ошибки, которая является нормой разницы между значением истинной координаты и предсказанием:

$$\varepsilon = \|y - f^W(x)\|.$$

Введем новую функцию правдоподобия для задачи локализации. Так как локализация является задачей регрессии, применим вероятностную модель из работы [25] для сглаженной  $L_1$  функции потерь и определим функцию правдоподобия как комбинацию функций правдоподобия Гаусса и Лапласа:

$$p(y|f^W(x), \sigma, \alpha) = \begin{cases} p_G(y|f^W(x), \sigma), & \varepsilon < \frac{1}{\beta^2}, \\ p_L(y|f^W(x), \alpha), & \text{иначе} \end{cases}$$

где  $p_G, p_L$  — функции правдоподобия Гаусса и Лапласа со скалярными значениями шума  $\sigma$  и  $\alpha$  соответственно.

Перейдем к максимизации логарифмической функции правдоподобия. Следуя выводу функции для задачи регрессии в случае  $L_2$  функции потерь [25], для сглаженной  $L_1$  функции ее можно записать как

$$\log p(y|f^W(x), \sigma, \alpha) \propto \begin{cases} -\frac{\varepsilon^2}{2\sigma^2} - \log \sigma, & \varepsilon < \frac{1}{\beta^2}, \\ -\alpha \varepsilon + \log \alpha, & \text{иначе} \end{cases},$$

где  $L_2$  — соответствует правдоподобию Гаусса  $p_G$ ;  $L_1$  — правдоподобию Лапласа  $p_L$ .

Получим целевую функцию минимизации:

$$L(W, \sigma, \alpha) = -\log p(y|f^W(x), \sigma, \alpha) \propto \begin{cases} \frac{1}{2\sigma^2} L_2(W) + \log \sigma, & \varepsilon < \frac{1}{\beta^2}, \\ \alpha L_1(W) - \log \alpha, & \text{иначе} \end{cases}, \quad (1)$$

где  $L_1(W)$  соответствует функции потерь  $L_1$ ,  $L_2(W) = \varepsilon^2$  соответствует евклидовой функции потерь.

Функция правдоподобия в (1) имеет два параметра отклонения  $\sigma, \alpha$ , соответствующие  $L_1$  и  $L_2$ . Однако для ее применения на реальных данных с неизвестными метками координат объектов это неудобно, так как истинные координаты ограничивающего прямоугольника и их разница с предсказанными координатами неизвестны. Чтобы найти зависимость между  $\sigma$  и  $\alpha$ , а также сохранить свойство распределения правдоподобия, из равенства функции плотности распределения единице получим следующий вид сглаженной  $L_1$  функции потерь:

$$B\text{Smooth}L_1(\varepsilon) = \begin{cases} \frac{\varepsilon^2}{2\sigma^2} + \log \sigma, & \varepsilon < \frac{1}{\beta^2} \\ -\beta^2 \varepsilon \log \tau + \log \tau + \frac{1}{2\sigma^2 \beta^4} + \log \sigma, & \text{иначе} \end{cases}.$$

Из равенства функции плотности распределения единице получим зависимость между дисперсиями:

$$\alpha = -\beta^2 \log \tau,$$

где  $\tau = 1 - \text{erf}\left(\frac{1}{\beta^2 \sqrt{2\sigma^2}}\right)$ ,  $\text{erf}$  — функция ошибки Гаусса [37].

Подробный вывод функции потерь приведен в работе [38]. На рис. 2 показаны графики предложенной байесовской и исходной сглаженной  $L_1$  функции потерь. Функция потерь штрафует нейронную сеть лучше, чем исходная: для данных, содержащих меньше шума, она

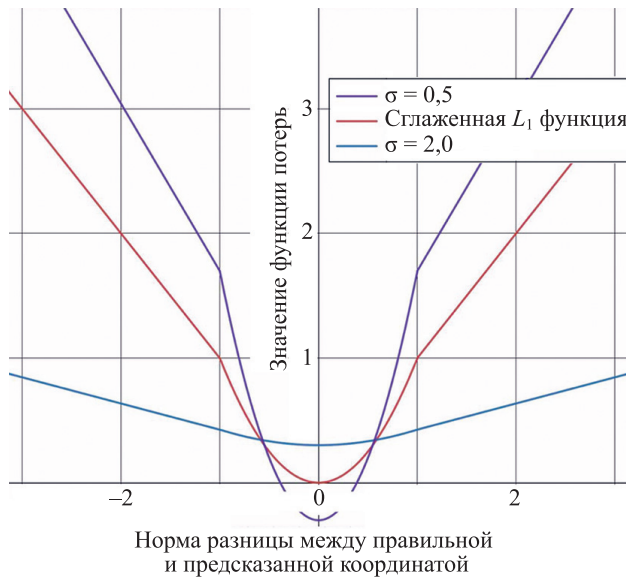


Рис. 2. Вид предлагаемой байесовской сглаженной функции потерь  $L_1$  при различных оценках алеаторной неопределенности (при  $\sigma = 0,5$ ,  $\sigma = 2,0$ ) в сравнении с исходной сглаженной функцией потерь  $L_1$  (красная линия)  
 Fig. 2. The proposed Bayesian Smooth  $L_1$  loss function for different estimates of aleatoric uncertainty at  $\sigma = 0.5$ ,  $\sigma = 2.0$  compared to the original Smooth  $L_1$  loss (red line)

больше штрафует нейронную сеть за большие ошибки прогнозирования. Для более шумных данных она штрафует нейронную сеть более равномерно, имея к разметке меньше доверия.

### Байесовская фокусная функция потерь

Введем функцию правдоподобия для задачи классификации, которая представляет собой модифицированную фокусную функцию потерь. В RetinaNet для задачи классификации использована логистическая функция активации, что более удобно для наборов данных с не исключаящими друг друга классами. По аналогии с [25] для задачи классификации функция правдоподобия может быть определена как:

$$p(y|f^W(x), \sigma) = \text{Sigmoid}\left(\frac{1}{\sigma^2}f^W(x)\right),$$

где коэффициент шума  $\sigma$  отражает гомоскедастичную неопределенность. Эта функция правдоподобия также может быть интерпретирована как распределение Больцмана, где величина  $f^W(x)$  масштабируется на  $\frac{1}{\sigma^2}$ .

Теперь вместо максимизации логарифма правдоподобия максимизируем фокусную функцию потерь, которая добавляет к логарифму правдоподобия множитель  $(1 - p)^y$  и сонаправлена логарифмическому правдоподобию. Для вывода BFL из исходной функции потерь FL, необходимо освободить  $f^W(x)$  в логистической функции правдоподобия от коэффициента масштабирования  $\frac{1}{\sigma^2}$ . Для решения этой задачи и получения новой формы правдоподобия использованы переходы, описанные в работе [38].

Конечный вид правдоподобия на основе фокусной функции потерь

$$BFL(p(y|f^W), \sigma) = -\left(\frac{1}{\sigma^2} \log p_t - \log \sigma\right) \frac{1}{\sigma} (1 - p_t)^{\sigma^2} y.$$

Байесовская фокусная функция потерь равна исходной функции при  $\sigma = 1$ , т. е. при  $p(y|f^W(x), \sigma) = \text{Sigmoid}(f^W(x))$ .

На рис. 3 показаны графики сравнения предложенной байесовской фокусной и исходной функций потерь. Предложенная функция потерь штрафует нейронную сеть лучше, чем исходная ( $\sigma = 1$ ): для менее зашумленных данных штрафует нейронную сеть меньше за хорошо классифицируемые объекты и больше — за большие ошибки предсказания. Для более шумных данных она более равномерно штрафует нейронную сеть, меньше доверяя разметке.

### Байесовская RetinaNet

Для многозадачной байесовской RetinaNet с выходом  $y_1$  для задачи локализации и  $y_2$  для задачи классификации получим целевую задачу минимизации:

$$L(f^W(x), y_1, y_2, \sigma_1, \sigma_2) = BSmoothL_1(f^W(x), y_1, \sigma_1) + \alpha BFL(f^W(x), y_2, \sigma_2),$$

где  $BSmoothL_1(f^W(x), y_1, \sigma_1)$  — байесовская сглаженная  $L_1$  функция потерь для  $y_1$ ;  $BFL(f^W(x), y_2, \sigma_2)$  — байесовская фокусная функция потерь для  $y_2$ ;  $\alpha$  — коэффициент балансировки, который регулирует влияние  $BFL(f^W(x), y_2, \sigma_2)$ .

Следуя [25], в настоящей работе сеть обучалась выводить логарифмическую дисперсию  $s = \log \sigma^2$ , что стабильнее численно, чем непосредственная регрессия дисперсии  $\sigma^2$ , так как позволяет избежать деления на ноль и взятия логарифма нуля.

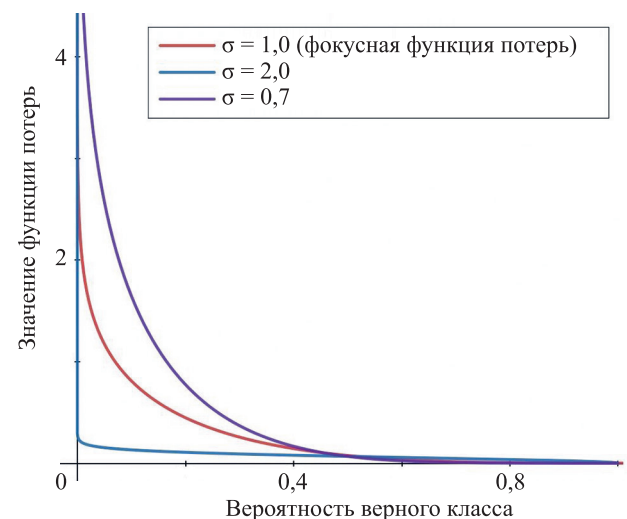


Рис. 3. График предложенной байесовской фокусной функции потерь при различных уровнях алеаторной неопределенности (при  $\sigma = 0,7$ ,  $\sigma = 1,0$ ,  $\sigma = 2,0$ ). При  $\sigma = 1,0$  функция равна исходной фокусной функции потерь  
 Fig. 3. The proposed Bayesian Focal loss function for different estimates of aleatoric uncertainty at  $\sigma = 0.7$ ,  $\sigma = 1.0$ ,  $\sigma = 2.0$ . At  $\sigma = 1.0$  the function is equal to the original Focal loss

Полученная многозадачная функция потерь  $L(f^W(x), y_1, y_2, \sigma_1, \sigma_2)$  минимизируется по отношению к  $W$ , а также  $\sigma_1$  и  $\sigma_2$ , отражающими гомоскедастичную неопределенность.

### Набор данных

Для экспериментов был использован набор данных изображений пыльцы, полученных с помощью оптического микроскопа<sup>1</sup>, созданный и размеченный автором данной работы. Набор данных включает в себя 13 основных видов растений-аллергенов, использованных в исследовании. Сложность распознавания этих видов растений определяется в первую очередь многообразием форм, размеров, направлений их пыльцевых зерен и фокусов микроскопа. Пыльцевые зерна разных родов растений часто имеют схожую округлую форму, а отличить пыльцевые зерна одного рода, но разных видов, иногда тяжело даже для опытного палинолога. В табл. 1 приведены примеры изображений исследуемых видов растений, которые были получены с помощью оптического микроскопа Olympus BX51 с системой визуализации Olympus DP71. Набор данных насчитывает 1100 изображений и всего 5100 пыльцевых зерен. Для экспериментов набор данных был разделен на train/test в пропорции 70/30.

### Результаты

В экспериментах в качестве сети для первичного извлечения признаков (backbone) для RetinaNet и байесовской RetinaNet использовалась ResNet-50 [39] ввиду ограничения доступной виртуальной памяти. Архитектура байесовской RetinaNet была такой же, как и оригинальная RetinaNet. Отличались лишь функции потерь, которые в байесовской RetinaNet были заменены на предложенные в настоящей работе. Для предварительной обработки изображений использовались масштабирование и нормализация со стандартизацией. Для рассмотренных моделей использован масштаб изображения, равный 800 пикселям.

Эксперименты выполнены на одном графическом процессоре NVIDIA GeForce GTX 1080 Ti с 10 ГБ VRAM. Оригинальная реализация модели RetinaNet взята из репозитория<sup>2</sup>, основанного на фреймворке keras [40].

Сначала оригинальная модель была обучена с помощью оптимизатора Adam [41] с изначальной скоростью обучения 0,00001 с применением уменьшения ее на плато. Использовались стандартные аугментации: зеркаливание, небольшие повороты, сдвиги.

Далее была обучена предложенная байесовская RetinaNet модель с тем же оптимизатором и гиперпараметрами, что и оригинальная.

Значения инициализации  $s_1 = \log \sigma_1^2$  для задачи локализации равно 1,  $s_1 = \log \sigma_2^2$  для задачи классификации — 0. Обучение обеих моделей заняло около 30 эпох, что в среднем составляет около одного часа.

Для каждой архитектуры было обучено пять моделей для формирования доверительных интервалов. В качестве меры сравнения использована стандартная для детекции мера средней точности (Average Precision), подсчитывающая среднюю точность детекции (AP) при разных порогах Intersection Over Union (IoU). AP рассчитывалась для каждого класса или вида растения отдельно, а затем усреднялась по всем классам, формируя mean Average Precision (mAP), также стандартную меру оценки в задаче детекции.

Табл. 2 представляет результаты сравнения метрик предложенной и исходной моделей RetinaNet-ResNet-50. Исходная модель достигла 93,56 % mAP на тестовом наборе данных, в то время как предложенная — 96,32 % mAP, превзойдя исходную.

В результате предложенная байесовская RetinaNet модель обеспечивает увеличение на 2,76 % по метрике mAP. Кроме того, дисперсия полученного результата меньше, чем у исходной модели, что свидетельствует о лучшей ее калибровке. Оценки алеаторной неопределенности, полученные в ходе обучения, составили 1,56 для задачи регрессии и 0,95 для задачи классификации. Эти значения согласуются с тем фактом, что набор данных обладает зашумленной разметкой. На самом деле в разметке набора данных присутствует большое количество пыльцевых зерен, снятых с расфокусировкой, которые не были размечены как пыльца. Вероятно, они вносят большую дисперсию относительно информации о местоположении пыльцы и без моделирования алеаторной неопределенности затрудняют обучение нейросети. Дисперсия классификационной разметки же близка к нормальному распределению. Так, на некоторых изображениях с пыльцевой ловушки были представлены пыльцевые зерна разных видов растений, которые могли быть размечены неверно в то время, как большинство видов растений являлись гербарными и однородными по видам в составах проб.

### Выводы

Полученный результат подтверждает утверждение о том, что моделирование гомоскедастичной алеаторной неопределенности позволяет повысить точность решения задачи детекции. Также можно сделать вывод о том, что предполагаемые функции потерь действительно штрафуют нейронную сеть лучше, чем оригинальные.

В будущем возможно применение предложенных функций потерь к другим моделям, которые основаны на архитектуре RetinaNet, например, SpineNet [42], ATSS [43]. Кроме того, интересно применение разработанных функций потерь для моделирования гетероскедастичной алеаторной неопределенности, что может больше повысить точность детекции и интерпретируемость детекции каждого объекта.

Разработанные функции потерь масштабируемы на другие наборы данных [38], и могут быть применены к детекторам, обучаемым с помощью перекрестной

<sup>1</sup> Набор данных пыльцы с оптического микроскопа [Электронный ресурс]. URL: [http://genome.ifmo.ru/files/papers\\_files/Allergy2018/](http://genome.ifmo.ru/files/papers_files/Allergy2018/) (дата обращения: 19.07.2021).

<sup>2</sup> Keras-RetinaNet [Электронный ресурс]. <https://github.com/fizyr/keras-retinanet> (дата обращения: 19.07.2021).

Таблица 1. Примеры изображений разных видов растений-аллергенов из исследуемого набора данных  
 Table 1. Examples of allergenic plants pollen images from the studied dataset






Виды аллергенов	Примеры изображений
Ольха клейкая	
Липа сердцевидная	
Злаки	
Береза повислая	
Ива белая	

Таблица 2. Сравнение модели RetinaNet, обученной с исходными функциями потерь, и байесовской RetinaNet, обученной с предлагаемыми функциями потерь, которые моделируют гомоскедастичную алеаторную неопределенность на тестовом множестве целевого набора данных пыльцевых зерен с оптического микроскопа

Table 2. Comparison of the RetinaNet trained with original loss functions and Bayesian RetinaNet trained with proposed loss functions, which model homoscedastic aleatoric uncertainty, on the test set of the studied pollen dataset taken with an optical microscope

Виды аллергенов	RetinaNet, AP, %	Байесовская RetinaNet, AP, %
Ива белая	97,57 ± 0,32	97,43 ± 0,56
Липа сердцевидная	97,61 ± 0,18	97,51 ± 0,50
Ольха клейкая	85,97 ± 2,74	90,32 ± 1,79
Береза повислая	83,67 ± 1,68	84,64 ± 1,13
Крапива	95,54 ± 1,19	96,16 ± 1,00
Маревые	93,66 ± 4,74	97,58 ± 1,33
Подорожник	86,48 ± 10,62	97,67 ± 0,65
Щавель	97,46 ± 1,15	97,34 ± 1,46
Злаки	100,00	100,00
Сосна	97,24 ± 1,31	99,17 ± 0,30
Клен	97,82 ± 0,53	98,32 ± 0,62
Лещина обыкновенная	98,36 ± 2,02	100,00
Полынь	84,83 ± 1,77	96,01 ± 0,97
Среднее	93,56 ± 1,45	96,32 ± 0,29

энтропии или фокусной функции потерь для задачи классификации, а также  $L_1$  или сглаженной  $L_1$  функции потерь для задачи локализации.

### Заключение

В работе предложены и применены новые функции потерь для задачи детекции изображений пыльцы, а именно, функции: *Байесовская фокусная функция потерь* и *Байесовская сглаженная функция потерь  $L_1$* . Предлагаемые функции способны моделировать гомо-

скедастичную алеаторную неопределенность во время обучения модели и не требуют изменений архитектуры.

Предлагаемые функции потерь исследованы на задаче детекции (локализации и классификации) пыльцевых зерен на изображениях с микроскопа на основе модели RetinaNet. В результате исследования достигнуто увеличение mAP на 2,76 %. Полученный результат подтверждает гипотезу о том, что моделирование гомоскедастичной алеаторной неопределенности повышает точность решения задачи детекции.

## Литература

## References

- Pawankar R., Canonica G.W., Holgate S., Lockey R. *WAO White Book on Allergy*. USA: World Allergy Organization, 2013. 242 p.
- Bousquet P.J., Chinn S., Janson C., Kogevinas M., Burney P., Jarvis D. Geographical variation in the prevalence of positive skin tests to environmental aeroallergens in the European Community Respiratory Health Survey I // *Allergy*, 2007. V. 62. N 3. P. 301–309. <https://doi.org/10.1111/j.1398-9995.2006.01293.x>
- D'Amato G., Cecchi L., Liccardi G. Thunderstorm-related asthma: not only grass pollen and spores // *Journal of Allergy and Clinical Immunology*, 2008. V. 121. N 2. P. 537–539. <https://doi.org/10.1016/j.jaci.2007.10.046>
- De Weger L.A., Bergmann K.C., Rantio-Lehtimäki A., Dahl A., Buters J., Déchamp C., Belmonte J., Thibaudon M., Cecchi L., Besancenot J.-P., Galán C., Waisel Y. Impact of Pollen // *Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts*. Springer, 2013. P. 161–215. [https://doi.org/10.1007/978-94-007-4881-1\\_6](https://doi.org/10.1007/978-94-007-4881-1_6)
- Jaeger S. The trouble with threshold values for allergy forecasts // *Aerobiological Monographs. Towards a comprehensive vision* / ed. by B. Clot, P. Comtois, B. Escamilla-Garcia. 2006. P. 233–245.
- Caillaud D.M., Martin S., Segala C., Besancenot J.-P., Clot B., Thibaudon M., Nonlinear short-term effects of airborne Poaceae levels on hay fever symptoms // *Journal of Allergy and Clinical Immunology*, 2012. V. 130. N 3. P. 812–814. <https://doi.org/10.1016/j.jaci.2012.04.034>
- Committee for Medicinal Products for Human Use. European Medicines Agency Committee for Medicinal Products for Human Use (CHMP) guideline on the evaluation of anticancer medicinal products in man. London, UK: European Medicines Agency, 2006.
- Sikoparija B., Skjøth C.A., Celenk S. et al. Spatial and temporal variations in airborne Ambrosia pollen in Europe // *Aerobiologia*, 2017. V. 33. N 2. P. 181–189. <https://doi.org/10.1007/s10453-016-9463-1>
- Novoselova L.V., Minaeva N. Pollen monitoring in Perm Krai (Russia)-experience of 6 years // *Acta Agrobotanica*, 2015. V. 68. N 4. P. 343–348. <https://doi.org/10.5586/aa.2015.042>
- Pfaar O., Bastl K., Berger U., Buters J., Calderon M.A., Clot B., Darsow U., Demoly P., Durham S.R., Gala'n C., Gehrig R., Gerth van Wijk R., Jacobsen L., Klimek L., Sofiev M., Thibaudon M., Bergmann K.C. Definition von Pollenexpositionszeiten für klinische Studien zur Allergen-Immuntherapie bei polleninduzierter Rhinokonjunktivitis—ein EAACI-Positionspapier // *Allergologie*, 2018. V. 41. N 9. P. 386–389. (in German). <https://doi.org/10.5414/ALX02053>
- Holt K.A., Bennett K.D. Principles and methods for automated palynology // *New Phytologist*, 2014. V. 203. N 3. P. 735–742. <https://doi.org/10.1111/nph.12848>
- Flenley J.R. The problem of pollen recognition // *Problems in Picture Interpretation* / ed. by M.B. Clowes, J.P. Penny. Canberra: CSIRO, 1968. P. 141–145.
- Boucher A., Hidalgo P.J., Thonnat M., Belmonte J., Galan C., Bonton P., Tomczak R. Development of a semi-automatic system for pollen recognition // *Aerobiologia*, 2002. V. 18. N 3-4. P. 195–201. <https://doi.org/10.1023/A:1021322813565>
- Chen C., Hendriks E.A., Duin R.P.W., Reiber J.H.C., Hiemstra P.S., de Weger L.A., Stoel B.C. Feasibility study on automated recognition of allergenic pollen: grass, birch and mugwort // *Aerobiologia*, 2006. V. 22. N 4. P. 275–284. <https://doi.org/10.1007/s10453-006-9040-0>
- Ronneberger O., Schultz E., Burkhardt H. Automated pollen recognition using 3D volume images from fluorescence microscopy // *Aerobiologia*, 2002. V. 18. N 2. P. 107–115. <https://doi.org/10.1023/A:1020623724584>
- Chica M. Authentication of bee pollen grains in bright-field microscopy by combining one-class classification techniques and image processing // *Microscopy Research and Technique*, 2012. V. 75. N 11. P. 1475–1485. <https://doi.org/10.1002/jemt.22091>
- Chudyk C., Castaneda H., Leger R., Yahiaoui I., Boochs F. Development of an automatic pollen classification system using shape, texture and aperture features // *CEUR Workshop Proceedings*, 2015. V. 1458. P. 65–74.
- Khanzhina N., Putin E. Pollen recognition for allergy and asthma management using gist features // *Communications in Computer and Information Science*, 2016. V. 674. P. 515–525. [https://doi.org/10.1007/978-3-319-49700-6\\_51](https://doi.org/10.1007/978-3-319-49700-6_51)
- Khanzhina N., Putin E., Filchenkov A., Zamyatina E. Pollen grain recognition using convolutional neural network // *Proc. 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2018. P. 409–414.
- Pawankar R., Canonica G.W., Holgate S., Lockey R. *WAO White Book on Allergy*. USA: World Allergy Organization, 2013, 242 p.
- Bousquet P.J., Chinn S., Janson C., Kogevinas M., Burney P., Jarvis D. Geographical variation in the prevalence of positive skin tests to environmental aeroallergens in the European Community Respiratory Health Survey I. *Allergy*, 2007, vol. 62, no. 3, pp. 301–309. <https://doi.org/10.1111/j.1398-9995.2006.01293.x>
- D'Amato G., Cecchi L., Liccardi G. Thunderstorm-related asthma: not only grass pollen and spores. *Journal of Allergy and Clinical Immunology*, 2008, vol. 121, no. 2, pp. 537–539. <https://doi.org/10.1016/j.jaci.2007.10.046>
- De Weger L.A., Bergmann K.C., Rantio-Lehtimäki A., Dahl A., Buters J., Déchamp C., Belmonte J., Thibaudon M., Cecchi L., Besancenot J.-P., Galán C., Waisel Y. Impact of Pollen. *Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts*. Springer, 2013, pp. 161–215. [https://doi.org/10.1007/978-94-007-4881-1\\_6](https://doi.org/10.1007/978-94-007-4881-1_6)
- Jaeger S. The trouble with threshold values for allergy forecasts. *Aerobiological Monographs. Towards a comprehensive vision*. Ed. by B. Clot, P. Comtois, B. Escamilla-Garcia, 2006, pp. 233–245.
- Caillaud D.M., Martin S., Segala C., Besancenot J.-P., Clot B., Thibaudon M., Nonlinear short-term effects of airborne Poaceae levels on hay fever symptoms. *Journal of Allergy and Clinical Immunology*, 2012, vol. 130, no. 3, pp. 812–814. <https://doi.org/10.1016/j.jaci.2012.04.034>
- Committee for Medicinal Products for Human Use. *European Medicines Agency Committee for Medicinal Products for Human Use (CHMP) guideline on the evaluation of anticancer medicinal products in man*. London, UK, European Medicines Agency, 2006.
- Sikoparija B., Skjøth C.A., Celenk S. et al. Spatial and temporal variations in airborne Ambrosia pollen in Europe. *Aerobiologia*, 2017, vol. 33, no. 2, pp. 181–189. <https://doi.org/10.1007/s10453-016-9463-1>
- Novoselova L.V., Minaeva N. Pollen monitoring in Perm Krai (Russia)-experience of 6 years. *Acta Agrobotanica*, 2015, vol. 68, no. 4, pp. 343–348. <https://doi.org/10.5586/aa.2015.042>
- Pfaar O., Bastl K., Berger U., Buters J., Calderon M.A., Clot B., Darsow U., Demoly P., Durham S.R., Gala'n C., Gehrig R., Gerth van Wijk R., Jacobsen L., Klimek L., Sofiev M., Thibaudon M., Bergmann K.C. Definition von Pollenexpositionszeiten für klinische Studien zur Allergen-Immuntherapie bei polleninduzierter Rhinokonjunktivitis—ein EAACI-Positionspapier. *Allergologie*, 2018, vol. 41, no. 9, pp. 386–389. (in German). <https://doi.org/10.5414/ALX02053>
- Holt K.A., Bennett K.D. Principles and methods for automated palynology. *New Phytologist*, 2014, vol. 203, no. 3, pp. 735–742. <https://doi.org/10.1111/nph.12848>
- Flenley J.R. The problem of pollen recognition. *Problems in Picture Interpretation*. Ed. by M.B. Clowes, J.P. Penny. Canberra, CSIRO, 1968, pp. 141–145.
- Boucher A., Hidalgo P.J., Thonnat M., Belmonte J., Galan C., Bonton P., Tomczak R. Development of a semi-automatic system for pollen recognition. *Aerobiologia*, 2002, vol. 18, no. 3-4, pp. 195–201. <https://doi.org/10.1023/A:1021322813565>
- Chen C., Hendriks E.A., Duin R.P.W., Reiber J.H.C., Hiemstra P.S., de Weger L.A., Stoel B.C. Feasibility study on automated recognition of allergenic pollen: grass, birch and mugwort. *Aerobiologia*, 2006, vol. 22, no. 4, pp. 275–284. <https://doi.org/10.1007/s10453-006-9040-0>
- Ronneberger O., Schultz E., Burkhardt H. Automated pollen recognition using 3D volume images from fluorescence microscopy. *Aerobiologia*, 2002, vol. 18, no. 2, pp. 107–115. <https://doi.org/10.1023/A:1020623724584>
- Chica M. Authentication of bee pollen grains in bright-field microscopy by combining one-class classification techniques and image processing. *Microscopy Research and Technique*, 2012, vol. 75, no. 11, pp. 1475–1485. <https://doi.org/10.1002/jemt.22091>
- Chudyk C., Castaneda H., Leger R., Yahiaoui I., Boochs F. Development of an automatic pollen classification system using shape, texture and aperture features. *CEUR Workshop Proceedings*, 2015, vol. 1458, pp. 65–74.
- Khanzhina N., Putin E. Pollen recognition for allergy and asthma management using gist features. *Communications in Computer and Information Science*, 2016, vol. 674, pp. 515–525. [https://doi.org/10.1007/978-3-319-49700-6\\_51](https://doi.org/10.1007/978-3-319-49700-6_51)
- Khanzhina N., Putin E., Filchenkov A., Zamyatina E. Pollen grain recognition using convolutional neural network. *Proc. 26th European*



20. Daood A., Ribeiro E., Bush M. Sequential recognition of pollen grain Z-stacks by combining CNN and RNN // Proc. 31<sup>st</sup> International Florida Artificial Intelligence Research Society Conference (FLAIRS). 2018. P. 8–13.
21. Sevillano V., Holt K., Aznarte J.L. Precise automatic classification of 46 different pollen types with convolutional neural networks // PLoS ONE. 2020. V. 15. N 6. P. e0229751. <https://doi.org/10.1371/journal.pone.0229751>
22. Schiele J., Rabe F., Schmitt M., Glaser M., Häring F., Brunner J.O., Bauer B., Schuller B., Traidl-Hoffmann C., Damialis A. Automated classification of airborne pollen using neural networks // Proc. 41<sup>st</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2019. P. 4474–4478. <https://doi.org/10.1109/EMBC.2019.8856910>
23. Northcutt C.G., Athalye A., Mueller J. Pervasive label errors in test sets destabilize machine learning benchmarks // arXiv.org. 2021. arXiv:2103.14749.
24. Kendall A., Gal Y. What uncertainties do we need in bayesian deep learning for computer vision? // Proc. 31<sup>st</sup> Annual Conference on Neural Information Processing Systems (NIPS). 2017. P. 5575–5585.
25. Cipolla R., Gal Y., Kendall A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics // Proc. 31<sup>st</sup> Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 7482–7491. <https://doi.org/10.1109/CVPR.2018.00781>
26. Bendale A., Boulton T.E. Towards open set deep networks // Proc. 29<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 1563–1572. <https://doi.org/10.1109/CVPR.2016.173>
27. Harakeh A., Smart M., Waslander S.L. BayesOD: A bayesian approach for uncertainty estimation in deep object detectors // Proc. 2020 IEEE International Conference on Robotics and Automation (ICRA). 2020. P. 87–93. <https://doi.org/10.1109/ICRA40945.2020.9196544>
28. Wirges S., Reith-Braun M., Lauer M., Stiller C. Capturing object detection uncertainty in multi-layer grid maps // Proc. 30<sup>th</sup> IEEE Intelligent Vehicles Symposium. 2019. P. 1520–1526. <https://doi.org/10.1109/IVS.2019.8814073>
29. Miller D., Nicholson L., Dayoub F., Sünderhauf N. Dropout sampling for robust object detection in open-set conditions // Proc. 2018 IEEE International Conference on Robotics and Automation (ICRA). 2018. P. 3243–3249. <https://doi.org/10.1109/ICRA.2018.8460700>
30. Miller D., Dayoub F., Milford M., Sünderhauf N. Evaluating merging strategies for sampling-based uncertainty techniques in object detection // Proc. 2019 International Conference on Robotics and Automation (ICRA). 2019. P. 2348–2354. <https://doi.org/10.1109/ICRA.2019.8793821>
31. Miller D., Sünderhauf N., Milford M., Dayoub F. Uncertainty for identifying open-set errors in visual object detection // arXiv.org. 2021. arXiv:2104.01328.
32. Postels J., Ferroni F., Coskun H., Navab N., Tombari F. Sampling-free epistemic uncertainty estimation using approximated variance propagation // Proc. 17<sup>th</sup> IEEE/CVF International Conference on Computer Vision (ICCV). 2019. P. 2931–2940. <https://doi.org/10.1109/ICCV.2019.00302>
33. Kraus F., Dietmayer K. Uncertainty estimation in one-stage object detection // Proc. 2019 IEEE Intelligent Transportation Systems Conference (ITSC). 2019. P. 53–60. <https://doi.org/10.1109/ITSC.2019.8917494>
34. Le M.T., Diehl F., Brunner T., Knol A. Uncertainty estimation for deep neural object detectors in safety-critical applications // Proc. 21<sup>st</sup> International Conference on Intelligent Transportation Systems (ITSC). 2018. P. 3873–3878. <https://doi.org/10.1109/ITSC.2018.8569637>
35. Lin T.-Y., Goyal P., Girshick R., He K., Dollar P. Focal loss for dense object detection // Proc. 16<sup>th</sup> IEEE International Conference on Computer Vision (ICCV). 2017. P. 2999–3007. <https://doi.org/10.1109/ICCV.2017.324>
36. Huber P.J. Robust estimation of a location parameter // *Breakthroughs in Statistics*. New York, NY: Springer, 1992. P. 492–518.
37. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables* / ed. by M. Abramowitz, I.A. Stegun, R.H. Romer. U.S. Government Printing Office, 1988. 1046 p.
38. Khanzhina N., Lapenok L., Filchenkov A. Towards robust object detection: Bayesian RetinaNet for homoscedastic aleatoric uncertainty modeling: preprint: submitted to the 37<sup>th</sup> Conference on Uncertainty in Artificial Intelligence (UAI 2021) [Электронный ресурс]. URL: [http://genome.ifmo.ru/files/papers\\_files/UAI](http://genome.ifmo.ru/files/papers_files/UAI) (дата обращения: 17.06.2021)
- Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2018, pp. 409–414.
20. Daood A., Ribeiro E., Bush M. Sequential recognition of pollen grain Z-stacks by combining CNN and RNN. *Proc. 31<sup>st</sup> International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 2018, pp. 8–13.
21. Sevillano V., Holt K., Aznarte J.L. Precise automatic classification of 46 different pollen types with convolutional neural networks. *PLoS ONE*, 2020, vol. 15, no. 6, pp. e0229751. <https://doi.org/10.1371/journal.pone.0229751>
22. Schiele J., Rabe F., Schmitt M., Glaser M., Häring F., Brunner J.O., Bauer B., Schuller B., Traidl-Hoffmann C., Damialis A. Automated classification of airborne pollen using neural networks. *Proc. 41<sup>st</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 4474–4478. <https://doi.org/10.1109/EMBC.2019.8856910>
23. Northcutt C.G., Athalye A., Mueller J. Pervasive label errors in test sets destabilize machine learning benchmarks. *arXiv.org*, 2021, arXiv:2103.14749.
24. Kendall A., Gal Y. What uncertainties do we need in bayesian deep learning for computer vision? *Proc. 31<sup>st</sup> Annual Conference on Neural Information Processing Systems (NIPS)*, 2017, pp. 5575–5585.
25. Cipolla R., Gal Y., Kendall A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. *Proc. 31<sup>st</sup> Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 7482–7491. <https://doi.org/10.1109/CVPR.2018.00781>
26. Bendale A., Boulton T.E. Towards open set deep networks. *Proc. 29<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1563–1572. <https://doi.org/10.1109/CVPR.2016.173>
27. Harakeh A., Smart M., Waslander S.L. BayesOD: A bayesian approach for uncertainty estimation in deep object detectors. *Proc. 2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 87–93. <https://doi.org/10.1109/ICRA40945.2020.9196544>
28. Wirges S., Reith-Braun M., Lauer M., Stiller C. Capturing object detection uncertainty in multi-layer grid maps. *Proc. 30<sup>th</sup> IEEE Intelligent Vehicles Symposium*, 2019, pp. 1520–1526. <https://doi.org/10.1109/IVS.2019.8814073>
29. Miller D., Nicholson L., Dayoub F., Sünderhauf N. Dropout sampling for robust object detection in open-set conditions. *Proc. 2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 3243–3249. <https://doi.org/10.1109/ICRA.2018.8460700>
30. Miller D., Dayoub F., Milford M., Sünderhauf N. Evaluating merging strategies for sampling-based uncertainty techniques in object detection. *Proc. 2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2348–2354. <https://doi.org/10.1109/ICRA.2019.8793821>
31. Miller D., Sünderhauf N., Milford M., Dayoub F. Uncertainty for identifying open-set errors in visual object detection. *arXiv.org*, 2021, arXiv:2104.01328.
32. Postels J., Ferroni F., Coskun H., Navab N., Tombari F. Sampling-free epistemic uncertainty estimation using approximated variance propagation. *Proc. 17<sup>th</sup> IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 2931–2940. <https://doi.org/10.1109/ICCV.2019.00302>
33. Kraus F., Dietmayer K. Uncertainty estimation in one-stage object detection. *Proc. 2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 53–60. <https://doi.org/10.1109/ITSC.2019.8917494>
34. Le M.T., Diehl F., Brunner T., Knol A. Uncertainty estimation for deep neural object detectors in safety-critical applications. *Proc. 21<sup>st</sup> International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 3873–3878. <https://doi.org/10.1109/ITSC.2018.8569637>
35. Lin T.-Y., Goyal P., Girshick R., He K., Dollar P. Focal loss for dense object detection. *Proc. 16<sup>th</sup> IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>
36. Huber P.J. Robust estimation of a location parameter. *Breakthroughs in Statistics*. New York, NY: Springer, 1992, pp. 492–518.
37. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Ed. by M. Abramowitz, I.A. Stegun, R.H. Romer. U.S. Government Printing Office, 1988, 1046 p.
38. Khanzhina N., Lapenok L., Filchenkov A. Towards robust object detection: Bayesian RetinaNet for homoscedastic aleatoric uncertainty modeling: preprint. Submitted to the 37<sup>th</sup> Conference on Uncertainty in Artificial Intelligence (UAI 2021). Available at: [http://genome.ifmo.ru/files/papers\\_files/UAI](http://genome.ifmo.ru/files/papers_files/UAI) (accessed: 17.06.2021)

39. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proc. 29<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
40. Chollet F. et al. Keras: The python deep learning library // Astrophysics Source Code Library. 2018. P. ascl: 1806.022.
41. Kingma D.P., Ba J.L. Adam: A method for stochastic optimization // Proc. 3<sup>rd</sup> International Conference on Learning Representations (ICLR). 2015.
42. Du X., Lin T.-Y., Jin P., Ghiasi G., Tan M., Cui Y., Le Q.V., Song X. SpineNet: Learning scale-permuted backbone for recognition and localization // Proc. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. P. 11589–11598. <https://doi.org/10.1109/CVPR42600.2020.01161>
43. Zhang S., Chi C., Yao Y., Lei Z., Li S.Z. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection // Proc. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. P. 9756–9765. <https://doi.org/10.1109/CVPR42600.2020.00978>
39. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proc. 29<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
40. Chollet F. et al. Keras: The python deep learning library. *Astrophysics Source Code Library*, 2018, P. ascl: 1806.022.
41. Kingma D.P., Ba J.L. Adam: A method for stochastic optimization. *Proc. 3<sup>rd</sup> International Conference on Learning Representations (ICLR)*, 2015.
42. Du X., Lin T.-Y., Jin P., Ghiasi G., Tan M., Cui Y., Le Q.V., Song X. SpineNet: Learning scale-permuted backbone for recognition and localization. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 11589–11598. <https://doi.org/10.1109/CVPR42600.2020.01161>
43. Zhang S., Chi C., Yao Y., Lei Z., Li S.Z. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. *Proc. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 9756–9765. <https://doi.org/10.1109/CVPR42600.2020.00978>

**Автор**

**Ханжина Наталья Евгеньевна** — программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0003-0795-642X>, [Nehanzhina@gmail.com](mailto:Nehanzhina@gmail.com)

**Author**

**Natalia E. Khanzhina** — Software Developer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0003-0795-642X>, [Nehanzhina@gmail.com](mailto:Nehanzhina@gmail.com)

*Статья поступила в редакцию 11.06.2021*  
*Одобрена после рецензирования 30.06.2021*  
*Принята к печати 04.08.2021*

*Received 11.06.2021*  
*Approved after reviewing 30.06.2021*  
*Accepted 04.08.2021*



Работа доступна по лицензии  
 Creative Commons  
 «Attribution-NonCommercial»