

doi: 10.17586/2226-1494-2021-21-6-951-961

УДК 004.85

Метод отключения путей по расписанию на основе семплирования Монте-Карло и непрерывной релаксации для оценки эпистемической неопределенности в задаче классификации 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 обучения нейронной сети NASNet с помощью нового метода отключения путей по расписанию на основе семплирования Монте-Карло и непрерывной релаксации для оценки эпистемической неопределенности в задаче классификации пылевых зерен на изображениях. Приведено описание разработанного метода и выполнено его сравнение с известными методами оценки эпистемической неопределенности. Актуальность работы состоит в том, что рассматриваемый пылевой набор данных крайне мал для задач компьютерного зрения, что порождает высокую эпистемическую неопределенность нейросетевых классификаторов и ведет к их переобучению. Разработанный метод позволяет получить оценку данного вида неопределенности без изменения архитектуры нейронной сети и обеспечить более точное решение задачи классификации пылицы. **Метод.** Предложенный метод позволяет оценить эпистемическую неопределенность нейронных сетей с многопутевыми ячейками на основе отключения путей с расписанием с использованием непрерывной релаксации для повышения точности и калибровки моделей распознавания изображений. Метод позволяет превратить произвольную многопутевую нейронную сеть в байесовскую путем ее семплирования на этапе предсказания методом Монте-Карло с разными масками отключения путей для оценки неопределенности. Произведено тестирование метода для задачи классификации на основе архитектуры NASNet. Преимущества метода продемонстрированы на задаче классификации изображений пылицы. **Основные результаты.** С помощью разработанного метода повышена точность классификации 13 видов пылицы растений-аллергенов на изображениях в среднем на 0,73 % по сравнению с базовой сетью NASNet до значения 98,34 % по F_1 мере. Также улучшена калибровка и уменьшена эпистемическая неопределенность модели в два раза по сравнению с ансамблем NASNet. Показано, что непрерывная релаксация параметра вероятности отключения путей в процессе обучения нейронной сети позволяет повысить точность решения задач и уменьшить эпистемическую неопределенность модели. **Практическая значимость.** Метод способствует значительному повышению точности классификации пылицы на изображениях, что имеет решающее значение для автоматизации распознавания пылицы в целом. Результаты работы дают возможность автоматизировать процесс аэропалеонтологического мониторинга и сократить время информирования больных поллинозами для предупреждения симптомов аллергии. Разработанный метод может быть применен для обучения нейронной сети для других задач компьютерного зрения на любых наборах изображений.

Ключевые слова

байесовское глубокое обучение, вариационный вывод, эпистемическая неопределенность, распознавание пылицы, распознавание изображений, оценка неопределенности

Благодарности

Автор выражает благодарность М.Ю. Каширину, М.С. Петухову, Н.В. Минаевой, Л.В. Новоселовой, Г.А. Заморину, Т.А. Полевой, А.А. Фильченкову, Е.Б. Замятиной, И.В. Харисовой, Ю. Пинаевой за неоценимую помощь в исследовании.

Ссылка для цитирования: Ханжина Н.Е. Метод отключения путей по расписанию на основе семплирования Монте-Карло и непрерывной релаксации для оценки эпистемической неопределенности в задаче классификации пылицы на изображениях // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21, № 6. С. 951–961. doi: 10.17586/2226-1494-2021-21-6-951-961

© Ханжина Н.Е., 2021

Monte Carlo Concrete DropPath for epistemic uncertainty estimation in pollen images classification

Natalia E. Khanzhina

ITMO University, Saint Petersburg, 197101, Russian Federation

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

Abstract

The paper presents the results of a new method for training the NASNet neural network called Monte Carlo Concrete DropPath for epistemic uncertainty estimation to classify pollen images. The developed method is compared with existing methods for epistemic uncertainty estimation. The method turns an arbitrary multipath neural network into a Bayesian one by sampling from the predictive distribution. The Monte Carlo method samples different masks of DropPath to estimate uncertainty. Moreover, the probability of DropPath is optimized using continuous relaxation. The proposed method was tested for the classification task on the state-of-the-art NASNet architecture. The method demonstrated advantages on the task of classifying pollen images. The classification accuracy increased for 13 pollen species of allergen plants by 0.73 % on average compared to the baseline NASNet, reaching 98.34 % by F1 measure. Furthermore, the method increased calibration and reduced the epistemic uncertainty of the model by two times compared to the NASNets ensemble. It is shown that continuous relaxation of the DropPath probability parameter increases the accuracy of problem solving and reduces the epistemic uncertainty of the model. These results contribute to the automation of aeropalynological monitoring to reduce the time of informing patients who suffer from pollinosis and hence to prevent allergy symptoms. The developed method can be applied to train a neural network for other computer vision tasks on any image dataset.

Keywords

Bayesian deep learning, variational inference, epistemic uncertainty, pollen recognition, image recognition, uncertainty estimation

Acknowledgements

The author would like to thank Maxim Kashirin, Maxim Petukhov, Natalia Minaeva, Larisa Novoselova, Georgiy Zamorin, Tatyana Polevaya, Andrey Filchenkov, Elena Zamyatina, Irina Kharisova and Yuliya Pinaeva for their great help and useful comments.

For citation: Khanzhina N.E. Monte Carlo Concrete DropPath for epistemic uncertainty estimation in pollen images classification. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 6, pp. 951–961 (in Russian). doi: 10.17586/2226-1494-2021-21-6-951-961

Введение

Поллиноз, или пыльцевая аллергия — заболевание, которому подвержено более 20 % населения Европы [1–3]. Для помощи больным в некоторых регионах производится сезонный пыльцевой мониторинг воздуха [4–11]. Однако существующая сеть станций мониторинга не способна обеспечить больных ежедневной информацией о наличии аллергенной пыльцы, что приводит к серьезным экономическим и социальным издержкам [12]. Более того, территориальные различия пыльцевого спектра затрудняют свободное передвижение больных между регионами [13]. Таким образом, требуется агрегация информации и глобальная оценка проблемы [11, 14].

Анализ пыльцы, который производится палинологами на станциях пыльцевого мониторинга, включает в себя определение видов растений по их пыльце с помощью микроскопа. Этот процесс крайне трудоемкий и занимает до четверти времени всего пыльцевого анализа, что не позволяет обеспечить больных поллинозом релевантной информацией [13, 15]. Решение этой проблемы может быть достигнуто путем автоматизации процесса распознавания, что позволит своевременно предупреждать больных о необходимости начала лечения симптомов.

Распознавание пыльцы с использованием оптического микроскопа можно решить с помощью распознавания изображений. С данной задачей эффективно справляются глубокие нейронные сети.

Постановка задачи

Задача автоматизации распознавания изображений пыльцы была впервые поставлена в 1968 году [16]. С этого момента исследователи по всему миру решают данную задачу с помощью различных методов компьютерного зрения и машинного обучения. На ранних этапах исследования многие ученые выделяли интерпретируемые признаки пыльцевых зерен, понятные палинологам, такие как форма, размеры, объем, периметр, текстурные признаки, апертуры [17–22]. И на основе признаков строят стандартные модели машинного обучения: метод опорных векторов, линейный дискриминантный анализ, случайный лес, нейросети, k -ближайших соседей и др.

Успех решения задачи распознавания в значительной степени связан с развитием программного и аппаратного обеспечений для глубокого обучения и компьютерного зрения [23], а также доступности GPU-серверов и открытых передовых моделей. Впервые применимость методов глубокого обучения к задаче распознавания пыльцы была исследована в работе [24]. Последующие исследования по этой теме целиком лежат в области глубокого обучения [25–33].

В работах [25–29] адаптированы существующие современные сверточные нейронные сети (СНС) для классификации пыльцы, такие как VGG-16 [30], AlexNet [23], DenseNet [31]. Авторами работ [24, 32, 33] построены свои собственные легковесные СНС с более низкой производительностью.

В некоторых работах использованы очень глубокие архитектуры, которые чрезмерны для этой проблемы и требуют большого количества данных для обучения. В то же время проблемой является отсутствие больших открытых наборов данных для обучения. Существующие открытые наборы данных [21, 34, 35] невелики и представляют разные домены, поэтому их объединение не тривиально.

В настоящей работе использован новый пыльцевой набор данных, который доступен онлайн и включает в себя более 5100 пыльцевых зерен 13 видов растений, — самый большой открытый набор изображений аллергенной пыльцы¹. Однако этого объема данных недостаточно для эффективного обучения глубоких нейронных сетей, что чревато переобучением моделей. Переобучение приводит к завышенной уверенности моделей на новых данных и ошибкам определения аллергенов. В связи с этим задача работы — разработка методов классификации изображений пыльцы при условии малого количества данных.

Для решения данной задачи предложены следующие подходы: использование генеративных нейронных сетей (ГНС) с целью генерации искусственных изображений для предобучения нейросетевого классификатора; метод оценки достоверности модели или неопределенности ее предсказаний.

Оценка неопределенности может освободить палинологов от кропотливой работы по ручной перепроверке предсказаний в тех случаях, когда модель уверена, и позволит сосредоточиться только на тех изображениях, где неопределенность высока.

Исследователи выделяют несколько типов неопределенности. Тип неопределенности, вызванный небольшим объемом обучающих данных или их нерепрезентативностью, называется эпистемической или неопределенностью модели [36].

Оценка эпистемической неопределенности применительно к задаче распознавания пыльцы не была изучена научным сообществом. Современные работы в области оценки эпистемической неопределенности глубоких нейронных сетей основаны на вариационном выводе, а именно на методе Монте-Карло (MC), известном как Monte Carlo Dropout (MC Dropout или MCDO) [37], который является частью байесовского глубокого обучения.

Изначально MC Dropout был предложен для регуляризации полносвязных нейронных сетей, но он плохо применим к СНС. Для СНС разработаны методы, такие как DropBlock [38], DropLayer [39] и DropFilter [40], которые превосходят MC Dropout с точки зрения точности и силы регуляризации модели. Следуя тому же направлению, в [41] предложен метод DropPath, который отбрасывает пути в сверточных многопутевых ячейках. Иногда данный метод называют DropLayer. Метод DropLayer был развит в работе [42] для обучения передовой модели NASNet. Авторы предложили отключать пути в ячейках с вероятностью, которая линейно

увеличивается во время обучения и с глубиной сети. Эта техника, которую назвали ScheduledDropPath, показала свою эффективность по сравнению с MC Dropout и DropPath для задачи классификации ImageNet.

Методы DropBlock, DropLayer и DropFilter были адаптированы для оценки эпистемической неопределенности на основе MC сэмплирования [43], но MC ScheduledDropPath ранее не применялся, в том числе для задачи классификации изображений пыльцы.

Чтобы заполнить этот пробел, для оценки неопределенности и повышения точности и калибровки моделей классификации пыльцы в настоящей работе предложен новый метод отключения путей по расписанию на основе MC сэмплирования и непрерывной релаксации (MC Concrete DropPath).

Метод использует DropPath для вариационного вывода на основе MC сэмплирования во время предсказания. Также новый метод развивает идею расписания DropPath, производя его нелинейно таким образом, что для каждой ячейки сети вероятность отключения пути вычисляется на основе непрерывной релаксации. Релаксация обеспечивается заменой распределения Бернулли вероятности отключения на распределение Concrete, которое использовано в работе [44] для метода MC Dropout.

NASNet ScheduledDropPath

ScheduledDropPath предложен авторами модели NASNet [42]. NASNet — очень глубокая нейронная сеть и одна из наиболее точных для задачи классификации изображений. Сеть содержит два типа ячеек: нормальные и редукционные. Каждая ячейка представляет собой многопутевый модуль, структура которого найдена с помощью нейроархитектурного поиска (Neural Architecture Search, NAS).

Авторы утверждают, что регуляризация NASNet с помощью стандартного MC Dropout снизила производительность модели. По этой причине они адаптировали метод регуляризации сети FractalNet под названием DropPath [41]. DropPath можно рассматривать как перенос идеи DropFilter [40] на уровень путей в многопутевых нейронных сетях. Однако исходный DropPath не позволял регулировать NASNet должным образом. Тогда авторы модифицировали DropPath и добавили отключение путей с линейным расписанием, увеличивающим вероятность отключения по мере обучения, а также с глубиной сети обучения, и назвали новую технику ScheduledDropPath (рис. 1).

На рис. 1 представлена нейросеть, состоящая из трех ячеек, подобных ячейкам NASNet. Каждая из ячеек содержит три сверточных слоя (1×1 conv, 3×3 conv, 5×5 conv) и одного субдискретизирующего или пулингового слоя (3×3 pool), представляющих пути. На каждом шаге обучения происходит отключение одного из слоев или путей в ячейке, помеченных красным. При этом вероятность отключения пути увеличивается с глубиной сети и с количеством шагов обучения, что условно обозначено как drop gate со стрелкой.

Concrete Dropout

Непрерывная релаксация успешно применена для техники регуляризации Dropout [44]. Вероятность отключения нейронов p обычно равна 0,5 как гипер-

¹ [Электронный ресурс]. Режим доступа: <https://www.kaggle.com/nataliakhanzhina/pollen201det> (дата обращения: 22.11.2021).

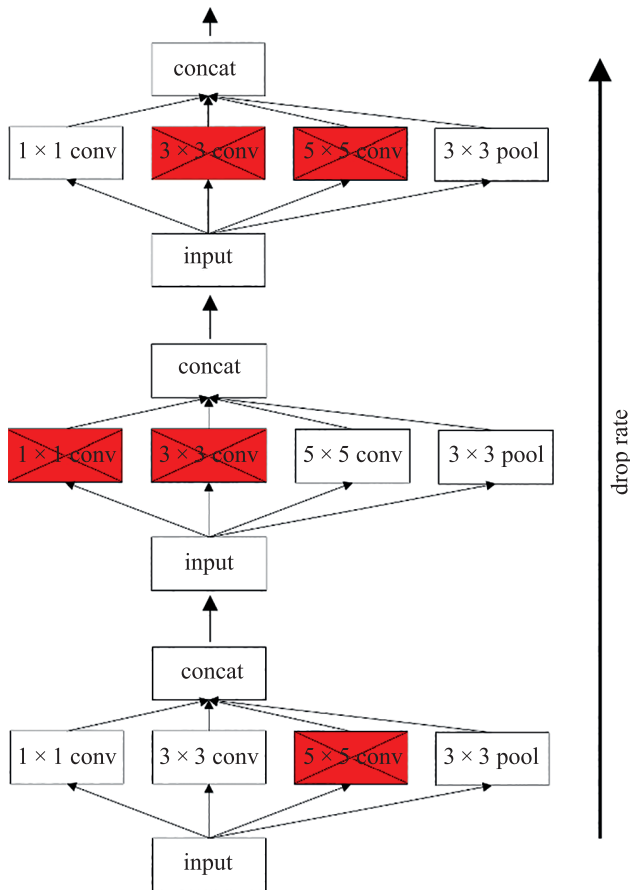


Рис. 1. Иллюстрация техники ScheduledDropPath для нейросети, состоящей из трех многопутевых сверточных ячеек

Fig. 1. An illustration of the ScheduledDropPath technique for a neural network consisting of three multipath convolutional cells

параметр модели и не меняется во время обучения, поскольку ее полный перебор требует больших вычислительных ресурсов. В [44] предложено принять p в качестве параметра для оптимизации. Чтобы сделать его дифференцируемым, дискретное распределение Бернулли вероятности p заменено на его непрерывную релаксацию с использованием распределения Concrete (Continuous+Discrete) [45].

Так как распределение Бернулли для случая Dropout одномерное, распределение Concrete сводится к логистическому распределению:

$$z = \text{sigmoid}\left(\frac{1}{t}(\log p - \log(1-p) + \log u - \log(1-u))\right), \quad (1)$$

где z — релаксация случайной величины; u — случайная величина, $u \sim \mathcal{U}(0, 1)$; t — температура.

Сигмоида подталкивает случайную величину z к 0 и 1.

Функция (1) дифференцируема по p , что позволяет применять алгоритм обратного распространения ошибки.

Метод Monte Carlo Concrete DropPath

В настоящей работе предложен способ трансформации произвольной многопутевой нейронной сети в байесовскую для оценки эпистемической неопре-

деленности. Это может быть достигнуто с помощью MC семплирования нейронной сети во время предсказания с разными масками ScheduledDropPath. Эксперименты показали, что метод MC ScheduledDropPath работает незначительно лучше, чем другие методы калибровки.

Таким образом, этот метод был усовершенствован с помощью нового подхода к расписанию DropPath, основанного на распределении Concrete [45]. Техника регуляризации названа Concrete DropPath, а соответствующий метод оценки неопределенности — MC Concrete DropPath (рис. 2).

Принцип работы нового метода MC Concrete DropPath состоит в применении непрерывной релаксации техники ScheduledDropPath на этапе предсказания моделью, который противопоставляется этапу обучения, для оценки неопределенности путем семплирования модели K раз с различными масками ScheduledDropPath (на рис. 2 приведен пример трех масок по направлению слева направо). В отличие от оригинального ScheduledDropPath, предложенный метод отключает пути в ячейках не с линейно растущей вероятностью, а с наиболее эффективной для каждой ячейки вероятностью, найденной с помощью релаксации распределения Бернулли распределением Concrete. Семплирование позволяет получить усредненное предсказание и оценку неопределенности.

Выполнено тестирование метода на NASNet, которое показало, что его можно применить к любой многопутевой нейронной сети, например: FractalNet [41], Inception-ResNet-v2 [46], ResNeXt [47] и др.

Целевая функция байесовской нейронной сети имеет вид

$$\mathcal{L}_{MC}(\theta) = -\frac{1}{M} \sum_{i \in S} \log p(y_i | f^\theta(x_i)) + \frac{1}{N} KL(q_\theta(\omega) || p(\omega)),$$

где $\log p(y_i | f^\theta(x_i))$ — функция правдоподобия модели; S — случайная выборка из M точек; $KL(q_\theta(\omega) || p(\omega))$ — дивергенция Кульбака–Лейблера (KL) между априорным распределением $p(\omega)$ и приближенным апостериорным распределением $q_\theta(\omega)$, представленным Scheduled/Concrete DropPath.

В то время как функция правдоподобия модели — обычная функция потерь нейронной сети (в рассматриваемом случае категориальная перекрестная энтропия), KL дивергенция представляет собой сумму KL дивергенций распределения параметров каждой ячейки НС. Для ячейки c KL определяется как:

$$KL(q_{M_c}(W) || p(W)) \propto \frac{(1-p_c)}{2s^2} \|M_c\|^2 - K\mathcal{H}(p_c),$$

где $\mathcal{H}(p_c)$ — энтропия вероятности p_c ; M_c — матрица средних весов ячейки; $q_{M_c}(W) = M_c^K D^K$, $D^K \sim \text{Bern}(1-p_c)$ — вектор вероятности отключения путей, K — количество путей в ячейке c ; s^2 — дисперсия априорного распределения.

Для метода MC ScheduledDropPath слагаемое $K\mathcal{H}(p)$ опускается, так как p зафиксирован. Но в случае Concrete DropPath p оптимизируется на основе замены распределения Бернулли распределением Concrete для

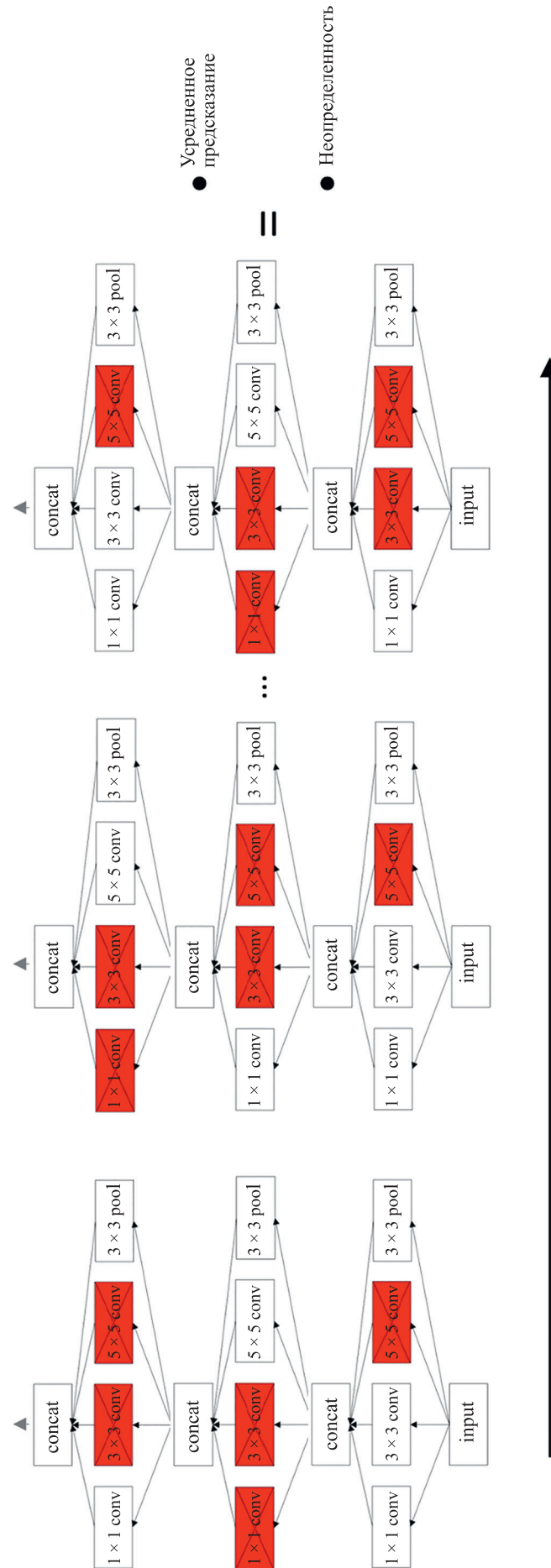


Рис. 2. Пример применения нового метода MC Concrete DropPath
 Fig. 2. The MC Concrete DropPath method illustration

каждой ячейки c подобно рассмотренному выше случаю Concrete Dropout. Следовательно, слагаемое $K\mathcal{H}(p)$ участвует в подсчете KL дивергенции.

Наконец, чтобы оценить эпистемическую неопределенность, необходимо сэмплировать аппроксимирующее распределение Concrete DropPath $q_{\theta}(\omega)$ N раз на этапе предсказания методом МС, сформировать вектор прогноза $\mathcal{Y} = \{y_1, \dots, y_N\}$. Далее \mathcal{Y} используется для расчета усредненного предсказания и оценки эпистемической неопределенности на основе меры взаимной информации (Mutual Information). Подробно предложенный метод описан в работе [48], который применен в настоящей работе для решения задачи классификации пыльцы.

StyleGAN

Модель NASNet — очень глубокая нейронная сеть, ее обучение на исходном наборе данных не давало должного обобщения модели, как и предварительное обучение на ImageNet. В связи с этим было решено произвести предобучение модели на искусственных изображениях пыльцы.

В настоящий момент наиболее эффективным методом генерации изображений являются ГНС (Generative Adversarial Network, GAN). В качестве ГНС в работе использована архитектура StyleGAN [49] — одна из наиболее передовых генеративных состязательных сетей для генерации изображений. Новизна и преимущество этой модели заключается в ее генераторе, который унаследовал генератор Progressive GAN [50], дополнив его адаптивной нормализацией экземпляров [51], отвечающую за передачу стиля.

Набор данных

Для экспериментов использован набор изображений пыльцы «POLLEN13L-det», полученных с помощью оптического микроскопа, созданный и размеченный автором данной работы. Набор данных¹ включает в

¹ [Электронный ресурс]. Режим доступа: <https://www.kaggle.com/nataliakhanzhina/pollen20ldet> (дата обращения: 22.11.2021).

себя 13 основных видов растений-аллергенов, характерных для России и насчитывает 5120 пыльцевых зерен. Гистограмма распределения по классам (видам растений) приведена на рис. 3.

Видно, что, некоторые классы представлены крайне малым числом изображений. Данное обстоятельство мотивирует необходимость применения ГНС для искусственного увеличения выборки, а также оценки эпистемической неопределенности моделей. Примеры изображений каждого вида растений приведены на рис. 4, а.

Эксперименты

Набор данных был разделен на тренировочную и тестовую выборки в соотношении 85/15 соответственно. Эксперименты проведены с помощью 5-кратной перекрестной валидации. В результате разделение на тренировочную, валидационную и тестовую выборки составило 68/17/15. Используются стандартные методы увеличения выборки (аугментации), такие как горизонтальные перевороты, сдвиги, вращения.

Каждая из сетей NASNet была предобучена на изображениях, полученных с помощью StyleGAN. В итоге сгенерировано 5000 изображений каждого класса, всего 65 000. Далее были выполнены следующие шаги обучения на основе оптимизатора Adam со скоростью обучения, равной 10^{-2} : модель была предварительно обучена на искусственных изображениях с косинусной функцией изменения скорости в течение 350 эпох; выполнено обучение переносом на реальные данные в течение 300 эпох с ранней остановкой; произведена тонкая настройка с косинусной функцией изменения скорости в течение 350 эпох.

Выполнено тестирование основанных на NASNet методов оценки неопределенности: NASNet MC Dropout (NASNet MCDO), NASNet MC DropFilter (NASNet MCDF), NASNet MC ScheduledDropPath (NASNet MCSDP, с предложенным методом), NASNet MC Concrete DropPath (NASNet MCCDP, с предложенным методом), глубокий ансамбль из пяти NASNet, обученных с помощью обычного ScheduledDropPath.

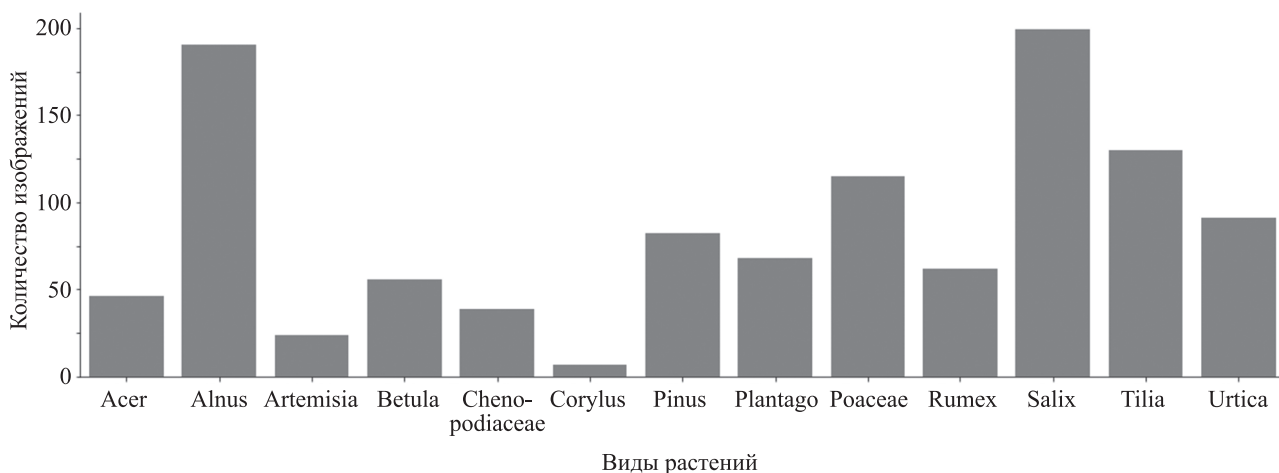


Рис. 3. Гистограмма распределения изображений по видам растений набора данных «POLLEN13L-det»

Fig. 3. Histogram of images per taxa of the “POLLEN13L-det” dataset

Наиболее важное сравнение — с глубоким ансамблем, так как это наиболее передовой метод калибровки моделей.

Для моделей MC Dropout, MC DropFilter, MC ScheduledDropPath использована вероятность отключения $p = 0,3$, которая обеспечила наилучшие результаты согласно экспериментам.

Для калибровки модели использованы четыре показателя: Negative Log Likelihood (NLL) [52], Brier Score (BS) [53], Expected Calibration Error (ECE), Maximum Calibration Error (MCE) [43, 54, 55]. Для оценки неопределенности была использована Mutual Information (MI) [36].

Результаты

Рассмотрим сравнение реальных изображений пыльцы с синтетическими, полученными с помощью архитектуры StyleGAN (рис. 4). Видно, что StyleGAN способна генерировать правдоподобные изображения пыльцы различных видов, сохраняя их специфическую текстуру, форму и цвет. Исключения составляют виды, представленные в наборе данных в крайне малом количестве (например, *Corylus*).

Модели, представленные в таблице, предварительно обучены на изображениях StyleGAN, а затем дообучены на реальных изображениях. Таблица демонстрирует результаты тестирования моделей на тестовой выборке. Результаты всех модели, кроме глубокого ансамбля, представлены с усреднением по кросс-валидации. Результаты MC моделей получены по 20 семплированиям, глубокого ансамбля — усреднением предсказаний по пяти моделям NASNet.

Показано, что модель NASNet MCCDP, обученная с предложенным методом, превосходит все остальные модели по точности и позволяет достичь 98,34 % по показателю F_1 меры. Применение непрерывной релаксации значительно повысило точность и калибровку модели. Кроме того, предложенный метод значительно лучше других моделей по метрикам калибровки: он обладает самым низким показателем Brier Score, Expected Calibration Error, при этом показатель Expected Calibration Error почти в два раза ниже, чем у глубокого ансамбля. Заметим, что метод обеспечил лучший показатель Mutual Information, т. е. оценки эпистемической неопределенности модели, более чем в два раза ниже, чем у глубокого ансамбля.

Выводы

В связи с тем, что используемый набор данных ранее был закрытым, а существующие открытые наборы данных относятся к другим доменам, произвести прямое сравнение результатов с другими работами невозможно.

Однако полученный результат превосходит все существующие результаты по задаче классификации, уступая лишь работам [23, 25], в которых использовались меньшие по величине наборы данных, а значит, получить на них высокую точность гораздо проще.

Отметим, что обучение и тестирование сети с MC DropFilter дало неожиданное падение в точности и калибровочных метриках модели по сравнению с MC Dropout. Также метод MC Scheduled DropPath показал довольно низкий результат. Таким образом, полученные результаты несколько расходятся с результатами тех же

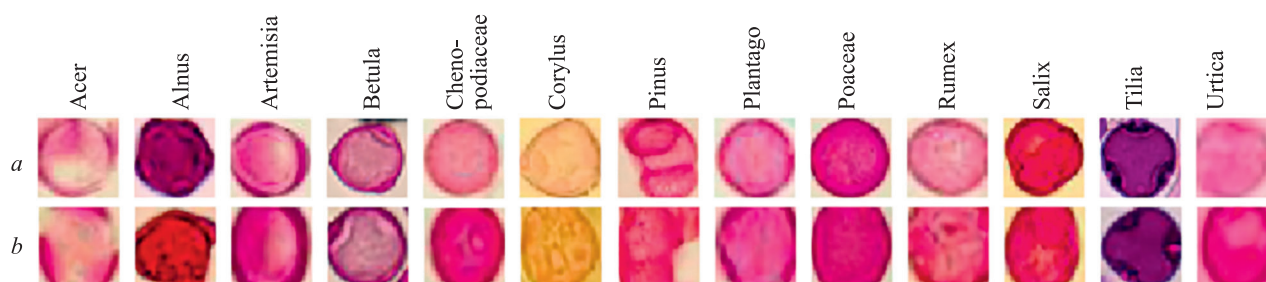


Рис. 4. Примеры изображений пыльцы: исходных (a) и созданных с помощью StyleGAN (b)

Fig. 4. Examples of real pollen images (a); images generated by StyleGAN (b)

Таблица. Сравнение точности и калибровки моделей NASNet, обученных с предложенным методом и другими существующими методами на тестовом множестве целевого набора данных пыльцевых зерен с оптического микроскопа
Table. Comparison of accuracy and calibration metrics for NASNet models trained with the proposed method and other existing methods on a test set of the target dataset of pollen images from lightning microscope

Модель	F_1 мера, %	$NLL \times 10^2$	$MI \times 10^2$	$BS \times 10^2$	$ECE \times 10^2$	MCE
NASNet	$97,608 \pm 0,485$	$9,607 \pm 0,996$	$6,427 \pm 1,760$	$4,366 \pm 0,402$	$4,282 \pm 1,550$	$2,271 \pm 0,519$
NASNet MCDO	$97,905 \pm 0,362$	$6,710 \pm 1,304$	$7,266 \pm 0,607$	$3,739 \pm 0,558$	$5,438 \pm 0,562$	$1,910 \pm 0,394$
NASNet MCDF	$88,613 \pm 0,553$	$26,214 \pm 0,627$	$36,134 \pm 0,380$	$17,020 \pm 0,324$	$27,611 \pm 0,363$	$2,659 \pm 0,312$
Глубокий ансамбль	97,59	8,07	7,24	4,58	4,88	2,079
NASNet MCSDP	$73,310 \pm 2,655$	$110,317 \pm 2,955$	$165,719 \pm 1,304$	$80,837 \pm 0,514$	$139,628 \pm 2,493$	$1,923 \pm 0,092$
NASNet MCCDP	$98,338 \pm 0,504$	$7,420 \pm 0,964$	$3,528 \pm 0,636$	$3,059 \pm 0,728$	$2,596 \pm 0,257$	$1,952 \pm 0,417$

методов на задаче распознавания опухолей мозга [48], что может быть вызвано спецификой предобучения на искусственных изображениях и задачи в целом.

В будущем планируется изучить предлагаемые методы на других многопутевых НС, таких как Inception-ResNet-v2 [46] и ResNeXt [47]. Concrete DropPath может быть эффективен не только для оценки неопределенности, но и в качестве техники регуляризации НС.

Заключение

В работе предложен и применен новый метод оценки эпистемической неопределенности для задачи

классификации изображений пыльцы, а именно, метод отключения путей по расписанию на основе семплирования Монте-Карло и непрерывной релаксации.

Метод исследован на задаче классификации (определении видов растений) пыльцевых зерен на изображениях, полученных с помощью оптического микроскопа на основе модели NASNet. В результате исследования достигнута точность классификации 98,34 % по F_1 мере для 13 видов растений и самые низкие значения по основным калибровочным метрикам, а также самая низкая оценка эпистемической неопределенности.

Литература

1. WAO White Book on Allergy / ed by R. Pawankar, G.W. Canonica, S.T. Holgate, R.F. Lockey, M.S. Blaiss. USA: World Allergy Organization, 2013.
2. 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>
3. Eigenmann P.A., Eigenmann P.A., Atanaskovic-Markovic M., O'B Hourihane J., Lack G., Lau S., Matricardi P.M., Muraro A., Namazova Baranova L., Nieto A., Papadopoulos N.G., Réthy L.A., Roberts G., Rudzeviciene O., Wahn U., Wickman M., Høst A. Testing children for allergies: why, how, who and when: an updated statement of the European Academy of Allergy and Clinical Immunology (EAACI) Section on Pediatrics and the EAACI-Clemens von Pirquet Foundation // *Pediatric Allergy and Immunology*, 2013. V. 24. N 2. P. 195–209. <https://doi.org/10.1111/pai.12066>
4. Bousquet J., Schunemann H.J., Fonseca J., Samolinski B. et al. MACVIA-ARIA Sentinel Network for allergic rhinitis (MASK-rhinitis): the new generation guideline implementation // *Allergy*, 2015. V. 70. N 11. P. 1372–1392. <https://doi.org/10.1111/all.12686>
5. Chikhladze M.V., Khachapuridze D.R., Sepiashvili R.I. The use of the Burkhard Pollen Trap to study the aerocological profile in Georgia // *Международный журнал по иммунореабилитации = International Journal on Immunorehabilitation*, 2009. T. 11. № 2. C. 200b–200b.
6. Kishikawa R., Sahashi N., Saitoh A., Kotoh E., Shimoda T., Shoji S., Akiyama K., Nishima S. Japanese cedar airborne pollen monitoring by Durham's and burkard samplers in Japan-estimation of the usefulness of Durham's sampler on Japanese cedar pollinosis // *Global Environmental Research*, 2009. V. 13. N 1. P. 55–62.
7. Puc M., Kotrych D., Lipiec A., Rapiejko P., Siergiejko G. Birch pollen grains without cytoplasmic content in the air of Szczecin and Białystok // *Alergoprofil*, 2016. V. 12. N 2. P. 101–105.
8. Piotrowska-Weryszko K., Weryszko-Chmielewska E., Rapiejko P., Rapiejko A., Malkiewicz M., Myszkowska D., Sulborska A., Żuraw B., Chłopek K., Voloshchuk K. Mugwort pollen season in southern Poland and Lviv (Ukraine) in 2015 // *Alergoprofil*, 2016. V. 12. N 1. P. 26–30.
9. Giorato M., Bordin A., Gemignani C., Turatello F., Marcer G. Airborne pollen in Padua (NE-Italy): A comparison between two pollen samplers // *Aerobiologia*, 2003. V. 19. N 2. P. 129–131. <https://doi.org/10.1023/A:1024405624565>
10. Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts / ed. by M. Sofiev, K.C. Bergmann. Springer, 2012. 252 p.
11. 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>
12. 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. 537–538. <https://doi.org/10.1016/j.jaci.2007.10.046>
13. Sikoparija B., Skjøth C.A., Celen S. et al. Spatial and temporal variations in airborne Ambrosia pollen in Europe // *Aerobiologia*,

References

1. WAO White Book on Allergy. Ed by R. Pawankar, G.W. Canonica, S.T. Holgate, R.F. Lockey, M.S. Blaiss. USA, World Allergy Organization, 2013.
2. 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>
3. Eigenmann P.A., Eigenmann P.A., Atanaskovic-Markovic M., O'B Hourihane J., Lack G., Lau S., Matricardi P.M., Muraro A., Namazova Baranova L., Nieto A., Papadopoulos N.G., Réthy L.A., Roberts G., Rudzeviciene O., Wahn U., Wickman M., Høst A. Testing children for allergies: why, how, who and when: an updated statement of the European Academy of Allergy and Clinical Immunology (EAACI) Section on Pediatrics and the EAACI-Clemens von Pirquet Foundation. *Pediatric Allergy and Immunology*, 2013, vol. 24, no. 2, pp. 195–209. <https://doi.org/10.1111/pai.12066>
4. Bousquet J., Schunemann H.J., Fonseca J., Samolinski B. et al. MACVIA-ARIA Sentinel Network for allergic rhinitis (MASK-rhinitis): the new generation guideline implementation. *Allergy*, 2015, vol. 70, no. 11, pp. 1372–1392. <https://doi.org/10.1111/all.12686>
5. Chikhladze M.V., Khachapuridze D.R., Sepiashvili R.I. The use of the Burkhard Pollen Trap to study the aerocological profile in Georgia. *International Journal on Immunorehabilitation*, 2009, vol. 11, no. 2, pp. 200b–200b.
6. Kishikawa R., Sahashi N., Saitoh A., Kotoh E., Shimoda T., Shoji S., Akiyama K., Nishima S. Japanese cedar airborne pollen monitoring by Durham's and burkard samplers in Japan-estimation of the usefulness of Durham's sampler on Japanese cedar pollinosis. *Global Environmental Research*, 2009, vol. 13, no. 1, pp. 55–62.
7. Puc M., Kotrych D., Lipiec A., Rapiejko P., Siergiejko G. Birch pollen grains without cytoplasmic content in the air of Szczecin and Białystok. *Alergoprofil*, 2016, vol. 12, no. 2, pp. 101–105.
8. Piotrowska-Weryszko K., Weryszko-Chmielewska E., Rapiejko P., Rapiejko A., Malkiewicz M., Myszkowska D., Sulborska A., Żuraw B., Chłopek K., Voloshchuk K. Mugwort pollen season in southern Poland and Lviv (Ukraine) in 2015. *Alergoprofil*, 2016, vol. 12, no. 1, pp. 26–30.
9. Giorato M., Bordin A., Gemignani C., Turatello F., Marcer G. Airborne pollen in Padua (NE-Italy): A comparison between two pollen samplers. *Aerobiologia*, 2003, vol. 19, no. 2, pp. 129–131. <https://doi.org/10.1023/A:1024405624565>
10. *Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts*. Ed. by M. Sofiev, K.C. Bergmann. Springer, 2012, 252 p.
11. 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>
12. 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. 537–538. <https://doi.org/10.1016/j.jaci.2007.10.046>
13. Sikoparija B., Skjøth C.A., Celen 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>

2017. V. 33. N 2. P. 181–189. <https://doi.org/10.1007/s10453-016-9463-1>
14. 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. <https://doi.org/10.5414/ALX02053>
 15. 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>
 16. Flenley J.R. The problem of pollen recognition // *Problems in Picture Interpretation* / ed. by M.B. Clowes, J.P. Penny. CSIRO: Canberra, 1968. P. 141–145.
 17. 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>
 18. 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>
 19. 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>
 20. 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>
 21. 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.
 22. 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
 23. Khanzhina N., Putin E., Filchenkov A., Zamyatina E. Pollen grain recognition using convolutional neural network // *Proc. of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*. 2018. P. 409–414.
 24. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // *Advances in Neural Information Processing Systems*. 2012. V. 2. P. 1097–1105.
 25. Daood A., Ribeiro E., Bush M. Classifying pollen using robust sequence alignment of sparse Z-stack volumes // *Lecture Notes in Computer Science*. 2016. V. 10072. P. 331–340. https://doi.org/10.1007/978-3-319-50835-1_31
 26. Sevillano V., Aznarte J.L. Improving classification of pollen grain images of the POLEN23E dataset through three different applications of deep learning convolutional neural networks // *PLoS ONE*. 2018. V. 13. N 9. P. e0201807. <https://doi.org/10.1371/journal.pone.0201807>
 27. 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>
 28. de Geus A.R., Barcelos C.A.Z., Batista M.A., da Silva S.F. Large-scale pollen recognition with deep learning // *Proc. of the 27th European Signal Processing Conference (EUSIPCO)*. 2019. P. 1–5. <https://doi.org/10.23919/EUSIPCO.2019.8902735>
 29. Gallardo-Caballero R., García-Orellana C.J., García-Manso A., González-Velasco H.M., Tormo-Molina R., Macías-Macías M. Precise pollen grain detection in bright field microscopy using deep learning techniques // *Sensors*. 2019. V. 19. N 16. P. 3583. <https://doi.org/10.3390/s19163583>
 30. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // *Proc. of the 3rd International Conference on Learning Representations, ICLR*. 2015.
 31. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks // *Proc. of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. P. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
 14. 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. <https://doi.org/10.5414/ALX02053>
 15. 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>
 16. Flenley J.R. The problem of pollen recognition. *Problems in Picture Interpretation*. Ed by M.B. Clowes, J.P. Penny. CSIRO, Canberra, 1968, pp. 141–145.
 17. 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>
 18. 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>
 19. 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>
 20. 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>
 21. 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.
 22. 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
 23. Khanzhina N., Putin E., Filchenkov A., Zamyatina E. Pollen grain recognition using convolutional neural network. *Proc. of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2018, pp. 409–414.
 24. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012, vol. 2, pp. 1097–1105.
 25. Daood A., Ribeiro E., Bush M. Classifying pollen using robust sequence alignment of sparse Z-stack volumes. *Lecture Notes in Computer Science*, 2016, vol. 10072, pp. 331–340. https://doi.org/10.1007/978-3-319-50835-1_31
 26. Sevillano V., Aznarte J.L. Improving classification of pollen grain images of the POLEN23E dataset through three different applications of deep learning convolutional neural networks. *PLoS ONE*, 2018, vol. 13, no. 9, pp. e0201807. <https://doi.org/10.1371/journal.pone.0201807>
 27. 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>
 28. de Geus A.R., Barcelos C.A.Z., Batista M.A., da Silva S.F. Large-scale pollen recognition with deep learning. *Proc. of the 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5. <https://doi.org/10.23919/EUSIPCO.2019.8902735>
 29. Gallardo-Caballero R., García-Orellana C.J., García-Manso A., González-Velasco H.M., Tormo-Molina R., Macías-Macías M. Precise pollen grain detection in bright field microscopy using deep learning techniques. *Sensors*, 2019, vol. 19, no. 16, pp. 3583. <https://doi.org/10.3390/s19163583>
 30. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proc. of the 3rd International Conference on Learning Representations, ICLR*, 2015.
 31. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks. *Proc. of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
 32. Menad H., Ben-Naoum F., Amine A. Deep convolutional neural network for pollen grains classification. *CEUR Workshop Proceedings*, 2019, vol. 2351.

32. Menad H., Ben-Naoum F., Amine A. Deep convolutional neural network for pollen grains classification // CEUR Workshop Proceedings. 2019. V. 2351.
33. 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. of the 41st 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>
34. Astolfi G., Gonçalves A.B., Menezes G.V., Borges F.S.B., Astolfi A.C.M.N., Matsubara E.T., Alvarez M., Pistori H. POLLEN73S: An image dataset for pollen grains classification // Ecological Informatics. 2020. V. 60. P. 101165. <https://doi.org/10.1016/j.ecoinf.2020.101165>
35. Gonçalves A.B., Souza J.S., Da Silva G.G., Cereda M.P., Pott A., Naka M.H., Pistori H. Feature extraction and machine learning for the classification of Brazilian Savannah pollen grains // PloS ONE. 2016. V. 11. N 6. P. e0157044. <https://doi.org/10.1371/journal.pone.0157044>
36. Kendall A., Gal Y. What uncertainties do we need in bayesian deep learning for computer vision? // Advances in Neural Information Processing Systems. 2017. V. 30. P. 5574–5584.
37. Gal Y., Ghahramani Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning // Proc. of the 33rd International Conference on Machine Learning (ICML). 2016. P. 1651–1660.
38. Ghiasi G., Lin T.Y., Le Q.V. DropBlock: A regularization method for convolutional networks // Advances in Neural Information Processing Systems. 2018. V. 31. P. 10727–10737.
39. Izmailov P., Podoprikin D., Garipov T., Vetrov D., Wilson A.G. Averaging weights leads to wider optima and better generalization // Proc. of the 34th Conference on Uncertainty in Artificial Intelligence (UAI). Vol. 2. 2018. P. 876–885.
40. Tompson J., Goroshin R., Jain A., LeCun Y., Bregler C. Efficient object localization using convolutional networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 648–656. <https://doi.org/10.1109/CVPR.2015.7298664>
41. Larsson G., Maire M., Shakhnarovich G. Fractalnet: Ultra-deep neural networks without residuals // Proc. of the 5th International Conference on Learning Representations (ICLR). 2017.
42. Zoph B., Vasudevan V., Shlens J., Le Q.V. Learning transferable architectures for scalable image recognition // Proc. of the 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 8697–8710. <https://doi.org/10.1109/CVPR.2018.00907>
43. Zhang Z., Dalca A.V., Sabuncu M.R. Confidence calibration for convolutional neural networks using structured dropout // arXiv.org. 2019. arXiv:1906.09551.
44. Gal Y., Hron J., Kendall A. Concrete dropout // Advances in Neural Information Processing Systems. 2017. V. 30. P. 3582–3591.
45. Maddison C.J., Mnih A., Teh Y.W. The concrete distribution: A continuous relaxation of discrete random variables // Proc. of the 5th International Conference on Learning Representations (ICLR). 2017.
46. Szegedy C., Ioffe S., Vanhoucke V., Alemi A.A. Inception-v4, inception-resnet and the impact of residual connections on learning // Proc. of the 31st AAAI Conference on Artificial Intelligence. 2017. P. 4278–4284.
47. Xie S., Girshick R., Dollár P., Tu Z., He K. Aggregated residual transformations for deep neural networks // Proc. of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 5987–5995. <https://doi.org/10.1109/CVPR.2017.634>
48. Khanzhina N., Kashirin M., Filchenkov A. Monte Carlo concrete DropPath for epistemic uncertainty estimation in brain tumor segmentation // Lecture Notes in Computer Science. 2021. V. 12959. P. 64–74. https://doi.org/10.1007/978-3-030-87735-4_7
49. Karras T., Laine S., Aila T. A style-based generator architecture for generative adversarial networks // Proc. of the 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 4396–4405. <https://doi.org/10.1109/CVPR.2019.00453>
50. Karras T., Aila T., Laine S., Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. // Proc. of the 6th International Conference on Learning Representations (ICLR). 2018.
51. Huang X., Belongie S. Arbitrary style transfer in real-time with adaptive instance normalization // Proc. of the 16th IEEE International Conference on Computer Vision (ICCV). 2017. P. 1510–1519. <https://doi.org/10.1109/ICCV.2017.167>
52. Loquercio A., Segu M., Scaramuzza D. A general framework for uncertainty estimation in deep learning. *IEEE Robotics and*

52. Loquercio A., Segu M., Scaramuzza D. A general framework for uncertainty estimation in deep learning // IEEE Robotics and Automation Letters. 2020. V. 5. N 2. P. 3153–3160. <https://doi.org/10.1109/LRA.2020.2974682>
 53. Ovadia Y., Fertig E., Ren J., Nado Z., Sculley D., Nowozin S., Dillon J.V., Lakshminarayanan B., Snoek J. Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift // Advances in Neural Information Processing Systems. 2019. V. 32.
 54. Naeini M.P., Cooper G., Hauskrecht M. Obtaining well calibrated probabilities using bayesian binning // Proc. of the 29th AAAI Conference on Artificial Intelligence. 2015. P. 2901–2907.
 55. Ashukha A., Lyzhov A., Molchanov D., Vetrov D. Pitfalls of in-domain uncertainty estimation and ensembling in deep learning // arXiv.org. 2020. arXiv:2002.06470.
56. Ashukha A., Lyzhov A., Molchanov D., Vetrov D. Pitfalls of in-domain uncertainty estimation and ensembling in deep learning // *Automation Letters*, 2020, vol. 5, no. 2, pp. 3153–3160. <https://doi.org/10.1109/LRA.2020.2974682>
 57. Ovadia Y., Fertig E., Ren J., Nado Z., Sculley D., Nowozin S., Dillon J.V., Lakshminarayanan B., Snoek J. Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift. *Advances in Neural Information Processing Systems*, 2019, vol. 32.
 58. Naeini M.P., Cooper G., Hauskrecht M. Obtaining well calibrated probabilities using bayesian binning. *Proc. of the 29th AAAI Conference on Artificial Intelligence*, 2015, pp. 2901–2907.
 59. Ashukha A., Lyzhov A., Molchanov D., Vetrov D. Pitfalls of in-domain uncertainty estimation and ensembling in deep learning. *arXiv.org*, 2020, arXiv:2002.06470.

Автор

Ханжина Наталья Евгеньевна — программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57192385166](https://orcid.org/0000-0003-0795-642X), <https://orcid.org/0000-0003-0795-642X>, Nehanzhina@gmail.com

Статья поступила в редакцию 21.10.2021
Одобрена после рецензирования 08.11.2021
Принята к печати 30.11.2021

Author

Natalia E. Khanzhina — Software Developer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57192385166](https://orcid.org/0000-0003-0795-642X), <https://orcid.org/0000-0003-0795-642X>, Nehanzhina@gmail.com

Received 21.10.2021
Approved after reviewing 08.11.2021
Accepted 30.11.2021



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»