

**DECISIONET З ПРОПОРЦІЙНИМ РОЗПОДІЛОМ ОБЧИСЛЮВАЛЬНИХ
ПОТУЖНОСТЕЙ / DECISIONET WITH PROPORTIONAL ALLOCATION OF
COMPUTATIONAL POWER**

Мокрий М. В., Швай Н. О. / Mokryi M., Shvai N.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-
Mohyla
Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: m.mokryy@ukma.edu.ua, n.shvay@ukma.edu.ua

This research focuses on DecisioNet (DN) — a Neural Decision Tree model that combines the generalization capabilities of deep networks with the computational efficiency of decision trees via conditional computation. The tree structure of the standard DN model allocates uniform computational power regardless of the number of classes per node. This study introduces Unbalanced DecisioNet (Unbalanced DN), which allocates computational power proportionally to the number of classes assigned to a node using hierarchical clustering. The results are compared with the original DN and baseline Network in Network models on CIFAR-10 and FashionMNIST datasets. Validation performed on CIFAR-10 demonstrated that Unbalanced DN improves classification accuracy compared to DN, which confirms the effectiveness of allocating increased computational resources to complex clusters while reducing resources for simple clusters.

У цьому дослідженні розглянуто нейронну мережу з бінарною деревоподібною структурою DecisioNet (DN) [1], яка належить до категорії нейронних дерев рішень [2] та представлено нову версію моделі з пропорційним розподілом обчислювальних ресурсів.

Традиційні нейронні мережі обробляють кожний вхідний зразок через усі параметри моделі, тоді як модель DN використовує умовні обчислення. Кожен вхідний зразок проходить по одному шляху від кореневого вузла до одного з листових вузлів дерева, активуючи лише частину від загальної кількості параметрів під час виведення. Класи призначаються вузлам на основі ієрархічної кластеризації за матрицею плутанини попередньо навченої базової моделі. Модель DecisioNet використовує архітектуру нейронної мережі Network in Network (NiN) [3].

Під час навчання використовується комбінована функція втрат, яка складається з функції втрат класифікації L_{cls} та маршрутизації L_{σ} , в якій параметр β контролює ступінь впливу над маршрутизацією вхідного зразка по дереву:

$$L = L_{cls} + \beta \cdot L_{\sigma}$$

У процесі навчання застосовується метод покращеного семантичного хешування [4], який поєднує диференційовану та дискретну маршрутизацію у рівних пропорціях. Це забезпечує зворотне поширення під час навчання, зберігаючи при цьому ефективність виведення завдяки вибору одного маршруту від кореня дерева до його листового вузла.

Структура збалансованого дерева моделі DN рівномірно розподіляє обчислювальні ресурси незалежно від розподілу класів між вузлами дерева. Тому з метою підвищення ефективності використання обчислювальних ресурсів було запропоновано підхід, за якого вони розподіляються пропорційно до обсягу даних, що обробляються на кожному вузлі.

Оновлена версія DN — UnbalancedDecisioNet (Unbalanced DN), реалізує нерівномірний розподіл обчислювальних ресурсів, дозволяючи виділяти більше ресурсів для складніших підзадач класифікації. Розподіл обчислювальних ресурсів для вузла i визначається за допомогою формули:

$$C_i = C_l \cdot 2^l \cdot \frac{n_i}{N}$$

Де N — загальна кількість класів, n_i — кількість класів призначених вузлу i та C_l — базова кількість параметрів для рівня дерева l у конфігурації DN.

Валідація створеної моделі була проведена на наборах даних FashionMNIST та CIFAR-10. Було натреновано базову модель NiN, а також моделі DN2 та Unbalanced DN2, де індекс «2» позначає глибину дерева.

Застосований метод ієрархічної кластеризації демонструє логічну структуру розподілу даних. Для набору даних CIFAR-10 початковий розподіл на першому рівні дерева ділить дані на два кластери: тварин та транспортні засоби. Для FashionMNIST перший рівень розподілу відокремлює взуття від одягу. Структуру дерева моделі Unbalanced DN2 для набору даних CIFAR-10, а також розподіл класів і обчислювальних ресурсів відповідно до ієрархічної кластеризації, наведено на рисунку 1.

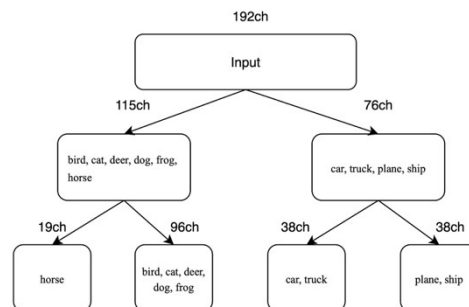


Рисунок 1. Структура моделі Unbalanced DN2 для CIFAR-10

Загальні результати експериментів наведено в таблиці 1. Вони демонструють, що моделі DecisioNet досягають подібної до базової моделі NiN результативності, використовуючи меншу кількість параметрів під час навчання, і ще меншу — під час виведення. Модель Unbalanced DN2 продемонструвала покращення точності класифікації на наборі даних CIFAR-10.

Таблиця 1. Загальні результати класифікації

Модель	Набір даних	Точність класифікації, %	К-сть параметрів
Baseline NiN	CIFAR-10	88.39	966,986
DN2	CIFAR-10	85.97	736,907
Unbalanced DN2	CIFAR-10	87.18	739,519
Baseline NiN	FashionMNIST	92.40	957,386
DN2	FashionMNIST	92.80	727,307
Unbalanced DN2	FashionMNIST	92.81	820,891

Емпірична валідація результатів моделі Unbalanced DN була проведена на наборах даних CIFAR-10 та FashionMNIST. Вона продемонструвала, що залежно від характеристик набору даних, складні вузли показують покращення точності класифікації за умов збільшення обчислювальних ресурсів, тоді як прості вузли зберігають високу результативність навіть при їх зменшенні.

Робота демонструє, що архітектура нейронного дерева рішень може включати нерівномірний розподіл обчислювальних ресурсів, відповідно до складності задачі класифікації, виходячи за межі рівномірного розподілу ресурсів по дереву рішень. У подальших дослідженнях планується перевірити ефективність моделі Unbalanced DN на більш складних наборах даних, зокрема CIFAR-100, а також дослідити вплив збільшення глибини архітектури DN на її ефективність.

Список джерел:

- Li H., Song J., Xue M., Zhang H., Ye J., Cheng L., Song M. A survey of neural trees // arXiv preprint arXiv:2209.03415. — 2022.
- Gottlieb N., Werman M. DecisioNet: A binary-tree structured neural network // Proceedings of the Asian conference on computer vision. — 2022. — P. 1675-1690.

3. Lin M., Chen Q., Yan S. Network in network // arXiv preprint arXiv:1312.4400. — 2013.
4. Kaiser Ł., Bengio S. Discrete autoencoders for sequence models // arXiv preprint arXiv:1801.09797. — 2018.

**МЕТОДИКА ТРЕНУВАННЯ ГЛИБОКОЇ НЕЙРОННОЇ МЕРЕЖІ DEEPLABV3
ДЛЯ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ ОБ'ЄКТІВ НА СУПУТНИКОВИХ
ЗОБРАЖЕННЯХ / METHODOLOGY FOR TRAINING DEEPLABV3 DEEP
NEURAL NETWORK FOR SEMANTIC SEGMENTATION OF OBJECTS IN
SATELLITE IMAGES**

Колінько П.В./ Kolinko P.V.

Національний університет “Києво-Могилянська Академія”/ National university Kyiv-
Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, каф. мультимедійних систем

E-mail: pavlo.kolinko@ukma.edu.ua

Semantic segmentation of satellite imagery remains one of the most challenging tasks in computer vision, requiring accurate pixel-level classification of diverse objects such as vegetation, buildings, and terrain features. DeepLab v3 is a state-of-the-art deep learning architecture with about 60-70M parameters to capture multi-scale contextual information. In this research, we proposed and tested a comprehensive training methodology for DeepLab integrated with PointRend head, utilizing ResNet-101 as backbone with layer-based learning rates and a gradual layer unfreezing strategy for the whole backbone. A composite loss function that is composed by weighted class-dependent loss metrics. Proposed approach allows to achieve the optimal results, prevent model overfitting and accelerate model convergence.

Семантична сегментація є фундаментальною та найобширнішою задачею комп'ютерного зору, що полягає у класифікації кожного пікселя зображення відповідно до його класу, на виході формуючи так звану “маску”. Застосування семантичної сегментації до супутникових знімків супроводжується специфічними викликами: надвисока роздільність зображень, велика варіативність масштабів об'єктів (для дерев - як поодинокі дерева, так і лісові масиви, для водойм - невеликі затоки або широкі озера), атмосферні перешкоди, сезонні зміни забарвлення рослинності, та значний дисбаланс класів. Проте, семантична сегментація супутникових зображень має дуже широке застосування. Вона охоплює моніторинг сільськогосподарських культур та оцінку врожайності, лісових ресурсів, міське планування (розпізнавання забудови) та аналіз змін землекористування, екологічний моніторинг водойм та прибережних зон, тощо. Успішне вирішення цієї задачі вимагає не лише потужних мереж глибокого навчання, але й спеціалізованих методів тренування, що враховують унікальні характеристики супутникових даних та специфіку цільових об'єктів сегментації.

Для вирішення такої задачі семантичної сегментації використовують великі згорткові нейронні мережі, розмірність яких нерідко доходить до 90+M параметрів, в залежності від архітектури. Прикладом такої глибокої нейронної мережі є Deeplabv3 [1]. Вона складається з таких ключових компонентів: CNN backbone, такий як ResNet з atrous convolutions, та модуля Atrous Spatial Pyramid Pooling (ASPP). Основною перевагою DeepLab над традиційними encoder-decoder архітектурами (як FCN або SegNet) є збереження високої роздільності feature maps завдяки атрозним згорткам, що дозволяє уникнути втрати просторової інформації. Порівняно з архітектурами на базі трансформерів, таких як наприклад Segment anything model (SAM), DeepLab [2] демонструє меншу обчислювальну складність та вимоги до пам'яті завдяки локальній природі згорткових операцій.

Використання backbone мережі ResNet-101 з попереднім навчанням на масштабному датасеті ImageNet (1.2 мільйона зображень, 1000 класів) забезпечує