

DOI: <https://doi.org/10.33216/1998-7927-2020-263-7-5-17>

УДК 004.8/93

## АНАЛІЗ МОЖЛИВОСТЕЙ КАПСУЛЬНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИРІШЕННЯ ПРОБЛЕМ, ПОВ'ЯЗАНИХ З РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

Арсенюк В.С., Барбарук В.М.

## ANALYSIS OF CAPSULE NEURAL NETWORKS POSSIBILITIES FOR SOLVING IMAGE RECOGNITION PROBLEMS

Arsenyuk V.S., Barbaruk V.M.

*У статті виконано огляд сьогочасних публікацій з досліджень щодо модифікацій базових алгоритмів маршрутизації, базової архітектури капсульної нейромережі та використання даними мережами датасетів. На підставі аналізу та узагальнення зроблені висновки щодо принципів побудови таких мереж на сучасному етапі, вибору алгоритму маршрутизації та оптимізатора для складних наборів даних.*

*Розглянуто питання щодо можливості застосування капсульного підходу при побудові нейронної мережі з розпізнавання зображень для мобільних додатків.*

**Ключові слова:** розпізнавання зображень, підвищення точності визначення об'єктів, капсульна нейронна мережа, мобільний додаток, полегшені моделі

**Вступ.** Всі передові мережі, що використовуються при вирішенні проблем, пов'язаних з розпізнаванням зображень, можуть розглядатися як різні варіації згортальної нейронної мережі (англ. Convolution Neural Networks, CNN).

Однак, погана здатність даних мереж розпізнавати позу, текстуру, обертання, деформації зображення або його частин, їх трансляційна інваріантність, наявність «проблеми Пікассо», призвела до впровадження капсульних нейронних мереж (англ. Capsule neural network, CapsNet).

**Постановка проблеми.** При використанні згортальних нейронних мереж в якості класифікатора, точність класифікації навченою моделлю може значно погіршуватися у разі зміни освітленості, кута обзору, або орієнтації реальних об'єктів.

Для вирішення цієї проблеми архітектуру CNN ускладнюють шляхом збільшення кількості згортальних шарів та карт ознак, що будуть враховувати масштаб, орієнтацію і таке інше. Але це призводить до значного зростання обсягу навчальних даних (для охоплення різних варіантів і уникнення перенавчання) та збільшення розміру переднавченої моделі. Все це зменшує можливість інтегрування таких мо-

делей в мобільні додатки. Крім того, такий підхід навчання моделі має тенденцію до запам'ятовування даних, а не до узагальнювання рішень.

Капсульний підхід здатний вирішити проблеми згортальних нейромереж та підвищити точність розпізнавання зображень, однак навчання та робота таких мереж потребує значних обчислювальних потужностей.

На сьогоднішній день безпосередньо мобільні пристрої, внаслідок обмеженої потужності апаратного забезпечення, не є платформою для моделей машинного навчання.

Дану проблему вирішують за допомогою створення переднавченої моделі, яку потім інтегрують в мобільний додаток.

Але й переднавчені моделі займатимуть на мобільних пристроях певний обсяг флеш-пам'яті і їх використання буде відбирати значні обчислювальні ресурси GPU і / або CPU, вимагаючи витрат енергії.

**Аналіз останніх досліджень і публікацій.** Після першої реалізації капсульної мережі [1], додаткових пояснень щодо можливостей трансформації архітектур глибокого навчання в прозорі нейромережі за допомогою включення капсул у різні їх шари [2], завдяки подальшим дослідженням та узагальненням, що сприяли розвитку капсульного підходу побудови нейромереж, з'явилась значна кількість модифікацій базової реалізації CapsNet, спрямованих на підвищення продуктивності даної мережі та усуненню ряду її обмежень. Для цього пропонувалися модифікації алгоритму маршрутизації, зміни в різних частинах базової архітектури капсульної нейронної мережі та інше.

Оскільки базові процедури маршрутизації CapsNet (включаючи динамічну маршрутизацію [1] та ЕМ-маршрутизацію [3]) є «дорогими» в обчислювальному відношенні, то для підвищення їх ефективності в подальших публікаціях різними авторами

були запропоновані наступні модифіковані процедури базових алгоритмів маршрутизації.

Для досягнення більш високої продуктивності S. Zhang, W. Zhao, X. Wu, Q. Zhou [4] пропонують використовувати капсули та динамічну маршрутизацію як ефективні доповнення до згорткових одиниць. Ними було запропоновано два спрощених алгоритми маршрутизації: швидку маршрутизацію на основі середнього зсуву (fast routing based on mean shift – *FRMS*) та швидку маршрутизацію на основі EM (fast routing based on EM – *FREM*), заснованих на непараметричній кластеризації та оцінці щільності ядра (kernel density estimation (KDE)). *FRMS* та *FREM* працюють швидше порівняно з алгоритмом EM-маршрутизації при деякій втраті точності на датасетах smallNORB, MNIST, Fashion-MNIST, CIFAR10 [4].

J. Choi, H. Seo, S. Im, M. Kang [5] замінюють динамічну маршрутизацію у базовій капсульній мережі на згорткове перетворення та «маршрутизацію уваги» – маршрутизацію між капсулами через модуль уваги. Під маршрутом уваги вони розуміють швидкий прохід вперед зі збереженням просторової інформації. Функцію активації *Squash* замінюють на функцію активації *tanh* на кожній капсулі каналу. Дані перетворення дають більш високу точність порівняно з базовою CapsNet, при використанні меншої кількості параметрів і меншого часу навчання [5].

W. Huang, F. Zhou для вдосконалення ієрархії капсул пропонують механізм подвійної уваги [6].

Для зменшення обмежень, викликаних обчислювальною складністю, H. Li, X. Guo, B. Dai та ін. апроксимують процес маршрутизації двома гілками: головною гілкою, що збирає первинну інформацію від її безпосереднього контакту в нижньому шарі, та допоміжною гілкою, яка поповнює головну на основі варіантів шаблонів, закодованих в інших нижчих капсулах. Ці дві гілки переплітаються своїми коефіцієнтами, так що шаблони функцій, закодовані в нижчих капсулах, можуть бути повністю використані та обмінені. Таке однопрохідне наближення є швидким, легким та контрольованим, порівняно з ітеративною, неконтрольованою схемою маршрутизації за домовленістю. Тому складність та час роботи моделі значно зменшуються [7].

H. Li, Y. Wang [8] підвищують ефективність розпізнавання капсульною мережею за рахунок оптимізації алгоритму маршрутизації і пропонують «алгоритм маршрутизації когнітивної послідовності капсульної мережі», що являє собою тримаршрутну маршрутизацію з пакетною нормалізацією [9] після кожного шару.

Z. Chen, D. Crandall [10] вбудовують процедуру маршрутизації в процедуру оптимізації з усіма іншими параметрами нейронної мережі, що дозволяє автоматизувати знаходження оптимальної кількості процедур маршрутизації.

D. Peer, S. Stabinger [11] експериментально показують, що маршрутизація за домовленістю не за-

безпечує автоматичного з'єднання капсули вищого рівня з декількома капсулами нижчого рівня для формування дерева синтаксичного аналізу. Для забезпечення формування дерева синтаксичного аналізу всіма активними капсулами вони пропонують алгоритм «динамічної глибокої маршрутизації» (dynamic deep routing), який дозволяє капсулам нижчого рівня вибрати єдину батьківську капсулу, а не відправляти свої вихідні дані в усі капсули більш високого рівня, як у випадку з вихідним алгоритмом маршрутизації за домовленістю [1]. Запропонований алгоритм дозволяє тренувати більш глибокі капсульні мережі, та є надійнішим для атак білих ящиків, ніж оригінальний алгоритм динамічної маршрутизації.

B. Jia, Q. Huang [12] наголошують на тому, що функція *Softmax* перетворює коефіцієнти зв'язку між капсулою верхнього шару і капсулою нижнього шару для алгоритму динамічної маршрутизації в насичений набір значень (щільно згрупований навколо початкового значення). Це може помилково відправити фонову інформацію до наступних шарів капсули із занадто великим коефіцієнтом, та призвести до неправильного підсумовування векторів прогнозування і тим чином вплинути на кінцевий результат класифікації. Щоб отримати більш рівномірний розподіл коефіцієнтів маршрутизації вони замінюють в алгоритмі функцію *Softmax* на функцію *Sigmoid*. Даний алгоритм отримав назву «алгоритму дисперсної динамічної маршрутизації» (disperse dynamic routing).

Зниження продуктивності мережі через цю проблему можна вирішити, використовуючи функцію *Max-Min* замість *Softmax* [13]. *Max-Min* виконує масштабно-інваріантну нормалізацію, дозволяючи капсулам нижнього рівня приймати незалежні значення на відміну від *Softmax*.

Алгоритм динамічної маршрутизації можна розглядати як задачу оптимізації, яку можна сформулювати як мінімізацію цільової функції. Він запобігає незбалансованості ймовірності активації із збільшенням кількості ітерацій [14]. Для стабілізування тренувального процесу, S. Sabour, N. Frosst, G. Hinton [1] виконали регуляцію матриці ваги, використовуючи втрати норми  $L_2$ . Більш загальне рішення може полягати в масштабуванні матриці ваги та забезпеченні того, щоб внутрішній добуток між вхідною та зваженою сумою усіх окремих прогнозів первинної капсули для капсули встановлювався нижче 1 для кожної ітерації [14].

Схема маршрутизації з однаковими вагами ініціалізації в базовій CapsNet має тенденцію до повільної збіжності і призводить до низької точності. Для покращення точності у класифікації з багатьма мітками на складних наборах даних S. Ramasinghe, C. D. Athuraliya, S. H. Khan пропонують моделювати початкові ваги маршрутизації як навчальні параметри, які можна навчити за допомогою зворотного поширення помилки. В задачах класифікації з кількома мітками прогнози первинних капсул CapsNet не є незалежними [15].

F. Ribeiro, G. Leontidis, S. Kollias [16] пропонують свій алгоритм маршрутизації капсул – «варіаційну байєсівську капсульну маршрутизацію» (variational bayes capsule routing або VB-Routing). Автори показують, що імовірнісний підхід забезпечує переваги перед базовими алгоритмами маршрутизації. Моделювання невизначеності параметрів капсули та ваг маршруту, дозволяє уникнути перенавчання та нестабільності тренування, що виникають в базових алгоритмах внаслідок припущення «єдиного батька» – за допомогою якого батьківська капсула може вимагати єдиного права опіки над дочірньою капсулою (точкою даних), що в свою чергу дає безкінечну ймовірність і нульову дисперсію. Байєсівський підхід усуває колапс дисперсії, шляхом моделювання невизначеності в параметрах пози капсули та дає конкурентоспроможні результати на даних CIFAR-10, Fashion-MNIST, SVHN, демонструє значне покращення розпізнавання MNIST [16].

R. LaLonde, U. Bagci модифікують оригінальний алгоритм динамічної маршрутизації, щоб діяти локально під час маршрутизації дочірніх капсул до батьківських капсул та обмінюватися матрицями перетворення між капсулами в межах одного типу капсул. Ці зміни різко зменшують пам'ять і навантаження на параметри оригінальної реалізації капсули, дозволяють працювати на великих розмірах зображення [17].

На відміну від вищезазначених підходів, зосереджених на підвищенні ефективності процедур маршрутизації, Z. Chen, X. Li, C. Wang, D. Crandall [18] намагаються їх взагалі видалити. Аргументуючи тим, що процедуру маршрутизації, призначену для отримання коефіцієнтів зв'язку між сусідніми шарами, можна вивчити та оптимізувати неявно, вони пропонують Pure CapsNets (P-CapsNets) без процедур маршрутизації. Маршрутизація за домовленістю розглядається як лінійні комбінації. Для зберігання капсул використовуються 3D-тензори. Запропонований алгоритм – це просто тензорне обчислення замість ітеративних процедур маршрутизації, яке значно пришвидшує навчання.

P-CapsNets може досягти кращої продуктивності, ніж багато інших варіантів CapsNets з різними процедурами маршрутизації, а також, ніж моделі глибокого стиснення, використовуючи значно менше параметрів. Однак P-CapsNets є більш вразливими до атак з використанням білого шуму порівняно з CNN [18]. Стаття знаходилася на розгляді для ICLR 2020 року, але в кінцевому підсумку була відкликана авторами.

За останні два роки різними авторами була запропонована значна кількість модифікацій базової архітектури [1] капсульної нейронної мережі.

З метою підвищення стабільності та швидкості збігу капсульної мережі було запропоновано ряд методів, що враховують інформацію щодо просторової структури. Mohammad Taha Bahadori [19] пропонує спектральну капсульну мережу (spectral capsule

network або S-Capsule network), що збігається швидше, ніж базова капсульна мережа з EM-маршрутизацією. Мережа вимірює збіг як ступінь вирівнювання голосів від капсул у нижчих шарах не в централізованому кластері, а в одновимірному лінійному підпросторі. Однак врахування тільки інформації про просторову структуру обмежує продуктивність капсульної мережі.

L. Zhang, M. Edraki, G. J. Qi [20] зосереджуються на формалізації принципової ідеї використання загальної довжини капсули, а не активації одного нейрона для моделювання присутності сутності [1, 3], вивчаючи групу капсульних підпросторів для представлення сукупності класів сутності. Вони стверджують, що ортогональна проекція на підпростір капсули грає вирішальну роль у забезпеченні її конкурентоспроможності, а просте групування нейронів в капсули, навпаки, не може поліпшити продуктивність. Проекцію капсули можна розглядати як багатовимірну нормалізацію ваги в підпросторах капсули, де звичайна нормалізація ваги – це просто приватний випадок проекції капсули на 1D лінії. Тому доцільно вивчити групу капсульних підпросторів, на які проектується вхідний вектор ознак. Потім довжини отриманих капсул використовувати для оцінки ймовірності належності до різних класів [20].

Через відсутність глибокої семантичної інформації базова CapsNet погано виконує завдання класифікації складних наборів даних. Для того, щоб отримати велике сприйнятливое поле в неглибокій згортковій структурі, в CapsNet використовується ряд великих згорткових ядер. Це збільшує кількість параметрів, що піддаються навчанню, і робить модель схильною до перенавчання. Для отримання надійних функцій та просторових взаємозв'язків із вхідних зображень C. Xiang, L. Zhang, Y. Tang та ін. [21] пропонують архітектуру MultiScale Capsule Network (MS-CapsNet), в якій спочатку структурна та семантична інформація отримується шляхом багатомасштабного вилучення ознак у згортковому шарі, а потім ієрархія об'єктів кодується у багатовимірну первинну капсулу у нижньому шарі оригінальної (базової) CapsNet. Використання багаторівневого згорткового ядра дозволяє зменшити кількість навчальних параметрів, а вдосконалення алгоритму відсіву для капсульного шару підвищує стійкість CapsNet.

A. Jaiswal, W. AbdAlmageed, Y. Wu, P. Natarajan [22] запропонували CapsuleGAN, що використовує CapsNet замість стандартних CNN як високопродуктивний дискримінатор в налаштуваннях генеративної змагальної мережі (generative adversarial network – GAN) під час моделювання зображень. CapsNets є кращими альтернативами CNN в якості дискримінаторів для GAN [23, 24], гарантуючи, що жодна важлива інформація не буде втрачена, як це може статися внаслідок операції об'єднання у CNN.

J. E. Lenssen, M. Fey, P. Libuschewski запропонували групові еквіваріантні шари капсул з векторами поз, обмежених як елементи групи. При зага-

льному алгоритмі маршрутизації за угодою та за певних умов, для таких груп можуть бути гарантовані еквіваріантність вихідних векторів поз а також інваріантність вихідних активацій [25]. Еквіваріантні капсульні шари вони з'єднують з груповими згортальними мережами, отримуючи архітектури глибоких нейронних мереж, що дозволяють виконувати розріджену оцінку оператора групової згортки, забезпечують контроль над конкретними властивостями еквівалентності та інваріантності і можуть використовувати маршрутизацію за угодою замість операцій об'єднання.

Експерименти N. H. Phong, B. Ribeiro показують, що CapsNet з додатковими конфігураціями, такими як збільшені кількості згортальних шарів після вхідного зображення та об'єднаних шарів реконструкції, на більш складних ніж MNIST та smallNORB наборах даних працюють зрівняно з моделями CNN, при цьому різко скорочуючи час навчання [26].

S. S. Phayre, A. Sikka, A. Dhall, D. Bathula вважають, що кращого вивчення карт характерних ознак для подальшого використання в CapsNets, згортальні шари можна замінити на щільно пов'язані згортки. Додавання прямих зв'язків між двома послідовними шарами допомагає вивчити кращі карти функцій, що, у свою чергу, допомагає у формуванні первинних капсул більш високої якості [27]. У запропонованій ними модифікації базової CapsNet – DCNet (Dense Capsule Network – щільна капсульна мережа), сформована більш глибока архітектура з багатшаровою щільною згортковою підмережею, що працює на основі пропусків з'єднань. Кожен шар з'єднується з наступним шаром прямим зв'язком, додаючись в остаточний шар згортки. Це призводить до кращого градієнтного потоку ніж у безпосередньо доданих згортальних шарів [27].

James O' Neill [28] представляє сіамські капсульні мережі, які можна використовувати для парних навчальних завдань. В даній мережі зкомбіновано попарні входи з капсульною архітектурою.

Y. Liu, Q. Zhang, D. Zhang, J. Han [29] пропонують двопотокову мережу виділення «частина-ціле» (Two-Stream Part-Object Assignment Network – TSPOANet) для виявлення об'єктів, помітних на деякому фоні (сегментування з фону).

Концепція глибокої згортково-деконволюційної архітектури мережі капсул, для передбачення міток об'єктів рівня пікселів, дозволяє компенсувати втрату глобальної інформації [17, 30].

Потреба класифікації складних наборів даних призводить до того, що архітектура капсульних мереж стає більш глибокою. До складу їх архітектур в якості «магістралей», заміників згортальних шарів включають повноцінні CNN, або їх значні частини [17, 22, 31–34]. Такі конфігурації значно скорочують перенавчання та підвищують точність розпізнавання.

Хоча капсульні нейромережі залишаються ще недостатньо дослідженими, вони вже почали використовуватися не тільки в експериментах з прости-

ми еталонними, а й з складними наборами даних. Додатки з CapsNets різноманітних архітектур та модифікацій алгоритмів маршрутизації знаходять застосування в різних соціально-економічних сферах: в астрономії для автоматизації обстежень великого неба та реконструкції кольорових зображень галактик датасету від Galaxy Zoo 2 [35], в метеорології для прогнозування екстремальних погодних явищ [36], в енергетиці у моделях багатоцільової оптимізації для бази даних з генерації відновлюваної енергії за допомогою вітру [37], в гірництві для виявлення підозрілих мікросейсмічних подій при підземних розробках [38], у біологічних дослідженнях [33], для класифікації структури білків з метою визначення їх функцій та допомоги в розробці ліків [39], у сільському господарстві для моніторингу росту рослин рису та запобігання пошкоджень їх хворобами й шкідниками за допомогою розпізнавання зображень рисових полів, отриманих від безпілотних літальних апаратів [40], при виявленні хвороб листя болгарського перцю за допомогою попередньо навченої на датасеті PlantVillage архітектурі [41], в екології для знаходження в реальному часі з повітря тварин та їх підрахунку [42], для класифікації сцен зйомок дистанційного зондування (запропонована архітектура оцінювалася на наборах даних дистанційного зондування зображень UC Merced Land-Use, AID та NWPU-RESISC45) [31].

За допомогою капсульних мереж вирішується задача пошуку дорожніх знаків для безпілотних транспортних засобів [43]. Запропонована мережа забезпечує стійкість до просторових дисперсій та досягає точності 97,6% на датасеті German Traffic Sign Recognition Benchmark (GTSRB).

Просторово-часова природа даних про дорожній рух, виражених в зображеннях, дозволяє застосовувати CapsNets для прогнозування потоку руху в складних дорожніх мережах [44]. Експерименти, проведені на однорічних даних, виміряних на відрізках доріг у місті Сантандер (Іспанія), демонструють, що запропонована CapsNet забезпечує більш точні прогнози швидкості транспорту, ніж CNN [44].

Per-Arne Andersen в своїй роботі демонструє життєздатність застосування капсульних нейромереж для ігрового штучного інтелекту. Поєднання CapsNet та глибокого навчання з підкріпленням дає можливість вдосконалити просунуті ігрові середовища [45].

D. Q. Nguyen, T. Vu, T. D. Nguyen та ін. [46] використовують капсульну мережу для моделювання трійки взаємозв'язків (суб'єкт, відношення, об'єкт), яка використовується для завершення графів знань та персоналізації пошуку. Запропонована ними CapsE перевершує інші моделі на базових наборах даних WN18RR та FB15k-237 для завершення графів знань та на SEARCH17 для персоналізації пошуку. Крім того, CapsE здатна ефективно моделювати відносини багато-до-багатьох [46].

Капсульні мережі знайшли застосування у сфері комп'ютерної лінгвістики. N. Zhang, S. Deng, Z. Sun та ін. [47] вважають, що CapsNet перетворюють проблему класифікації з кількома мітками в проблему багатобінарної класифікації та пропонують архітектуру капсульної мережі з вилучення відносин при автоматичній обробці текстів на природній мові для наборів даних NYT [48] (для вилучення парних сутностей) та Wikidata [49] (для вилучення відношень кількох пар сутностей). Результати, отримані ними показують, що модель покращує точність прогнозованих відносин [47].

CapsNets можуть працювати в якості структурного елемента розпізнавача мови жестів [50, 51]; визначати настрій і емоції виражені з урахуванням контексту слів [52], сигналів багатоканального електроенцефалографа [53, 54], за геометрією людського обличчя [55]; розпізнавати емоції мовлення, враховуючи просторові взаємозв'язки мовних особливостей у спектрограмах [56], виявляти наміри для систем «запитання-відповідь» [57], складати прогнози для рекомендаційних систем [58] і таке інше.

Для капсульних мереж з виявлення дій в якості вхідних даних можуть бути використані відео дані. Мережа VideoCapsuleNet [59] досягала найсучаснішої на той час продуктивності з виявлення багатьох дій, на наборах даних UCF-Sports, J-HMDB та UCF-101.

Представлена у [60] структура використовує капсули як для відео, так і для текстового подання. За допомогою візуально-текстової маршрутизації, мережа успішно сегментує дії у відео, відповідно до текстового запиту.

CapsNets успішно вирішують завдання комп'ютерної кримінології. В методі, запропонованому Н. Н. Nguyen, J. Yamagishi, I. Echizen, капсульна мережа застосовується для виявлення різного роду фальшивих зображень та відеопідробок (від атак відтворення з використанням друкованих зображень або записаних відео, до відеозаписів, створених за допомогою глибоких згорткових нейронних мереж) [34].

Генеративна змагальна мережа CapsuleGAN [22] з високопродуктивним, створеним з капсул дискримінатором, встановлює чи є зображення справжнім або штучно створеним (підробленим).

М. Н. Goldani, S. Momtazi, R. Safabakhsh використовують CapsNet в задачах виявлення фейкових новин [61], L. Bonan, Z. Tong, W. Min – у розпізнаванні відбитків пальців для біометричної ідентифікації [62]. Rui Ning застосовує капсульну мережу для виявлення дій зловмисних програм з виводу криптовалюти в браузері смартфонів [63].

Капсульні мережі розв'язують проблему виявлення певних звукових подій в оточуючому середовищі [64], класифікують аудіодані [65].

Однак найбільш перспективним напрямком застосування CapsNet все ж є робота із зображеннями. Мережі показують хороші результати при сегментації зображень [17, 66], розпізнаванні рукописних

символів [67], 3D-виявленні об'єктів [68], в задачах класифікації гіперспектральних зображень на реальних датасетах [69].

При аналізі медичних зображень дослідники стикаються з такими обмеженнями наборів даних як невелика кількість анотованих даних та дисбаланс класів (одна категорія даних може значно перевищувати інші), що є проблемою для ефективної роботи згортальних мереж [70]. Тому капсульні мережі, завдяки відсутності потреби у великих датасетах для навчання та відмінній спроможності щодо сегментації об'єктів, мають великий потенціал використання у медичній сфері.

Додатки, що використовують CapsNets, знайшли застосування в медичній діагностиці: для прогнозування хвороби Альцгеймера [71], класифікації типів пухлин головного мозку [72], апоптозу [72], вузликів на знімках комп'ютерної томографії при скринінгу раку легенів [74], діагностування пневмонії на рентгенівських знімках грудної клітини [75] та Covid-19 [76] за комп'ютерною томографією легень, ідентифікації раку грудей за зображеннями гістологічних препаратів [77], кількісної оцінки малярійної паразитемії [78], виявленні діабетичної ретинопатії [79], ішемічної хвороби серця [80] і таке інше.

Капсульні мережі можуть допомагати і в процесі лікування пацієнта. наприклад працювати у системах спостереження за пацієнтами, прикутими до інвалідних крісел або апаратів штучної вентиляції легенів, але зберігаючи свої когнітивні здатності [81].

**Мета статті.** За останні два роки було опубліковано значну кількість інформації щодо капсульного підходу побудови нейромереж. В статті проведено аналіз та узагальнення даних досліджень з огляду на можливість використання капсульних мереж як інструмента підвищення точності розпізнавання об'єктів для мобільних додатків.

**Результати досліджень.** В наступний час існують дві основні проблеми капсульних нейронних мереж з базовою архітектурою: їх обчислювальна складність та низька точність на складних датасетах.

Перша з яких з'явилася внаслідок того, що з використанням алгоритмів динамічної маршрутизації або ЕМ-маршрутизації, покликаних замінити операцію пулінгу і розширити можливість мережі реагувати на знаходження об'єкта в просторі, зріс час і складність обчислень при навчанні мережі. Замість простого вибору максимуму з деякої матриці використовуються ітеративні алгоритми, представлені в [1] та [3].

Складність механізму маршрутизації є основним фактором, що впливає на швидкість обчислень при навчанні. Крім нього, час навчання залежить від архітектури мережі (кількості шарів, кількості каналів і капсул в шарах, їх розмірності).

Кількість капсул зазвичай збільшується з ростом кількості класів класифікації, що призводить до ще більшого зростання великих обчислювальних ресурсів для навчання. Обсяги розрахунків залежать

також від складності і розміру вхідних даних (кількості змінних вектору або матриці нейронів, параметрів навчання, необхідних для роботи алгоритму маршрутизації).

Модель може реалізувати свій потенціал тільки при наявності ефективного способу підгонки моделі до даних [82]. Властивості набору даних мають вирішальне значення для продуктивності алгоритму. Алгоритм варто вибирати так, щоб він давав найкращий результат на обраному датасеті, а вже при цьому намагатися модифікувати швидкість його роботи.

Варіювання гіперпараметрів, таких як швидкість навчання, імпульс, розмір пакета, відсоток переривання та зниження швидкості навчання, не робить істотного впливу на продуктивність CapsNets. Однак відповідні операції маршрутизації і кількість ітерацій алгоритму маршрутизації є найбільш важливим гіперпараметром, що істотно впливає на продуктивність капсульної нейромережі [82].

N. Ohta, S. Kawai, H. Nobuhara [83] з метою зменшення обчислювальних витрат пропонують дозволяти одній капсулі представляти кілька класів об'єктів. Для визначення відстані між класами, вони включають метод дистанційного навчання ArcFace в функцію помилок.

H. Ren, J. Su, H. Lu [84] з метою зменшення кількості параметрів вводять механізм їх спільного використання (parameter-sharing). Параметри використовуються спільно шляхом сумісного використання матриці перетворення, що вибирає особливості капсул низького рівня для кожної капсули високого рівня. Кількість матриць перетворення в даному випадку буде дорівнювати кількості капсул високого рівня, а не добутку кількості капсул низького рівня на кількість капсул високого рівня.

Друга проблема пояснюється малою дослідженістю капсульного підходу до вирішення різних завдань.

Базова капсульна мережа добре показала себе на невеликих простих датасетах, та її продуктивність поки ще не може повноцінно суперничати з продуктивністю згортальних мереж на складних наборах даних, що мають високі внутрікласові варіації і фоновий шум, а тому потребує подальшого вивчення та вдосконалення. При цьому слід відзначити, що результати її роботи на складних датасетах порівняні з результатами згортальних мереж на такому ж етапі розвитку, що є одним з показників потенціалу капсульних нейромереж.

A. Chauhan, M. Babu, N. Kandru, S. Lokegaonkar довели, що збільшення кількості каналів в першому згортальному шарі, числа капсул в шарі первинних капсул, кількості капсул в наступних згортальних шарах призводить до більш швидкого навчання і збіжності CapsNets з алгоритмом динамічної маршрутизації та CapsNets з ЕМ-маршрутизацією. Збіжність дуже чутлива до кількості ітерацій маршрутизації. Оптимізатор *Adam* – кращий вибір для складних мереж, за ним йде *RMSProp* [85].

Точність класифікації капсульних продуктів залежить як від архітектури, так і від алгоритму маршрутизації [85].

До переднавчених моделей для мобільних пристроїв висувають наступні вимоги:

- обмеження на розмір моделей (оскільки їх потрібно завантажувати в оперативну пам'ять і віднімати значні обчислювальні ресурси GPU та / або CPU);

- обчислення, що виконує модель повинні бути досить ефективними для обробки значних обсягів інформації в розумний час, але без значного енергоспоживання або нагрівання батареї пристрою.

На це необхідно зважати при створенні мережі та використовувати так звані «полегшені» моделі, що будуть враховувати вищенаведені рекомендації.

Крім того, з метою зменшення розміру навченої моделі, необхідно оптимізувати модель для виводу, виконати операції «заморозки» та посттренинг квантування.

Це можна зробити за допомогою TensorFlow Lite, що використовує для навчання моделей бібліотеку TensorFlow.

TensorFlow генерує файл .pb – файл, що є серіалізованим графом обчислень через Protobuf. Цей файл не містить змінних. Змінні будуть зберігатися в .ckpt файлі.

Процес перетворення змінних в постійні (фіксовані значення) отримав назву заморозка. Заморожені графи (моделі) створюються за допомогою `freeze_graph.py`. Граф обчислень стає набагато простішим.

Опція `optimize_for_inference` видаляє непотрібні операції для виводу.

Квантування після навчання – це загальна методика, що дозволяє ще зменшити розмір моделі та в декілька разів зменшити затримку при невеликому погіршенні точності моделі [86].

Метод квантування з плаваючою комою включено в якості опції в конвертері моделей TensorFlow Lite.

**Висновок.** Хоча згортальні нейромережі є на сьогоднішній день одним з кращих засобів для розпізнавання зображень, та брак навчальних даних, трансляційна інваріантність CNN, їх схильність до змагальних нападів, можуть призвести до значного погіршення точності розпізнавання. Капсульний підхід побудови нейронних мереж дозволяє значно зменшити необхідну кількість навчальних даних, час навчання, параметрів системи у порівнянні з CNN. CapsNet можуть виявляти еквівалентні предмети, здійснювати реконструкцію зображень.

Капсульні нейромережі мають великий потенціал. З складними датасетами вони працюють порівняно зі згортальними мережами на такому ж етапі розвитку. Однак результати роботи CapsNet базової архітектури та базових алгоритмів маршрутизації на таких даних невисокі. Необхідність усунення ряду обмежень базової реалізації CapsNet та підвищення продуктивності мережі потребує її модифікації.

Точність класифікації капсульних продуктів залежить як від архітектури (кількості шарів, кількості каналів і капсул в шарах, їх розмірності), так і від алгоритму маршрутизації. Для підвищення точності розпізнавання та скорочення перенавчання при класифікації складних наборів даних доцільно робити архітектуру капсульної мережі більш глибокою, замінюючи згортальні шари повноцінними згортальними нейромережами, у яких відсутні пулінг-шари, або значними частинами CNN. Однак ускладнення архітектури впливає на час обчислень при навчанні. Властивості набору даних мають вирішальне значення для продуктивності алгоритму. Оскільки модель може реалізувати свій потенціал тільки при наявності ефективного способу підгонки моделі до даних, то алгоритм варто вибирати так, щоб він давав найкращий результат на обраному датасеті, а потім модифікувати швидкість його роботи. Моделі, що будуть використовуватися мобільними пристроями повинні бути полегшеними та оптимізованими.

### Література

1. Sabour S., Frosst N., G. Hinton. Dynamic routing between capsules. 2017. URL: <http://papers.nips.cc/paper/6975-dynamic-routing-between-capsules.pdf>.
2. Shahroudjeh A., Mohammadi A., Plataniotis K. N. Improved explainability of capsule networks: relevance path by agreement. 2018. URL: <https://arxiv.org/pdf/1802.10204.pdf>.
3. Hinton G., Sabou S., Frosst N. Matrix capsules with EM routing. 2018. URL: <https://openreview.net/pdf?id=HJWLFGWRb>.
4. Zhang S., Zhao W., Wu X., Zhou Q. Fast dynamic routing based on weighted kernel density estimation. 2018. URL: <https://arxiv.org/pdf/1805.10807.pdf>.
5. Choi J., Seo H., Im S., Kang M. Attention routing between capsules. 2019. URL: [https://openaccess.thecvf.com/content\\_ICCVW\\_2019/papers/NeurArch/Choi\\_Attention\\_Routing\\_Between\\_Capsules\\_ICCVW\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCVW_2019/papers/NeurArch/Choi_Attention_Routing_Between_Capsules_ICCVW_2019_paper.pdf).
6. Huang W., Zhou F. DA-CapsNet: dual attention mechanism capsule network. 2020. URL: <https://www.nature.com/articles/s41598-020-68453-w>.
7. Li H., Guo X., Dai B., Ouyang W., Wang X. Neural network encapsulation. 2018. URL: <https://arxiv.org/pdf/1808.03749.pdf>.
8. Li H., Wang Y. Cognitive consistency routing algorithm of capsule-network. 2018. URL: <https://arxiv.org/pdf/1808.09062.pdf>.
9. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015. URL: <https://arxiv.org/pdf/1502.03167.pdf>.
10. Chen Z., Crandall D. Generalized capsule networks with trainable routing procedure. 2018. — URL: <https://arxiv.org/pdf/1808.08692.pdf>.
11. Peer D., Stabinger S., Rodriguez-Sanchez A. Training deep capsule networks. 2018. URL: [https://www.researchgate.net/profile/Sebastian\\_Stabinger/publication/329945773\\_Training\\_Deep\\_Capsule\\_Networks/links/5c7e8901299bf1268d3a8928/Training-Deep-Capsule-Networks.pdf](https://www.researchgate.net/profile/Sebastian_Stabinger/publication/329945773_Training_Deep_Capsule_Networks/links/5c7e8901299bf1268d3a8928/Training-Deep-Capsule-Networks.pdf).
12. Jia B., Huang Q. DE-CapsNet: A diverse enhanced capsule network with disperse dynamic. 2020. URL: <https://www.mdpi.com/2076-3417/10/3/884>.
13. Zhao Z., Kleinhans A., Sandhu G., Patel I., Unnikrishnan K. P. Capsule networks with max-min normalization. 2019. URL: <https://arxiv.org/pdf/1903.09662.pdf>.
14. Wang D., Liu Q. An optimization view on dynamic routing between capsules. 2018. URL: <https://openreview.net/pdf?id=HJtFYJDf>.
15. Ramasinghe S., Athuraliya C. D., Khan S. H. A Context-aware Capsule Network for Multi-label Classification. 2018. URL: [https://openaccess.thecvf.com/content\\_ECCVW\\_2018/papers/11131/Ramasinghe\\_A\\_Context-aware\\_Capsule\\_Network\\_for\\_Multi-label\\_Classification\\_ECCVW\\_2018\\_paper.pdf](https://openaccess.thecvf.com/content_ECCVW_2018/papers/11131/Ramasinghe_A_Context-aware_Capsule_Network_for_Multi-label_Classification_ECCVW_2018_paper.pdf).
16. Ribeiro F., Leontidis G., Kollias S. Capsule routing via variational Bayes. 2019. URL: <https://arxiv.org/pdf/1905.11455.pdf>.
17. LaLonde R., Bagci U. Capsules for object segmentation. 2018. URL: <https://arxiv.org/abs/1804.04241>.
18. Chen Z., Li X., Wang C., Crandall D. Capsule networks without routing procedures. 2019. URL: <https://openreview.net/pdf?id=BlgNkrYvS>.
19. Bahadori M. T. Spectral capsule networks. 2018. URL: <https://openreview.net/pdf?id=HJuMvYPaM>.
20. Zhang L., Edraki M., Qi G. J. CapProNet: Deep feature learning via orthogonal projections onto capsule subspaces. 2018. URL: <http://papers.nips.cc/paper/7823-cappro-net-deep-feature-learning-via-orthogonal-projections-onto-capsule-subspaces.pdf>.
21. Xiang C., Zhang L., Tang Y., Zou W., Xu C. MS-CapsNet: A novel multi-scale capsule network. 2018. URL: [https://hal.archives-ouvertes.fr/hal-01908269/file/Xiang%20et%20al\\_2018\\_MS-CapsNet.pdf](https://hal.archives-ouvertes.fr/hal-01908269/file/Xiang%20et%20al_2018_MS-CapsNet.pdf).
22. Jaiswal A., AbdAlmageed W., Wu Y., Natarajan P. CapsuleGAN: Generative adversarial capsule network. 2018. URL: [https://openaccess.thecvf.com/content\\_ECCVW\\_2018/papers/11131/Jaiswal\\_CapsuleGAN\\_Generative\\_Adversarial\\_Capsule\\_Network\\_ECCVW\\_2018\\_paper.pdf](https://openaccess.thecvf.com/content_ECCVW_2018/papers/11131/Jaiswal_CapsuleGAN_Generative_Adversarial_Capsule_Network_ECCVW_2018_paper.pdf).
23. Saqr R., Vivona S. CapsGAN: Using dynamic routing for generative adversarial networks. 2018. URL: <https://arxiv.org/pdf/1806.03968.pdf>.
24. Majdabadi M. M., Ko S. B. MSG-CapsGAN: Multi-scale gradient capsule GAN for face super resolution. 2020. URL: <https://ieeexplore.ieee.org/abstract/document/9051244>.
25. Lenssen J. E., Fey M., Libuschevski P. Group equivariant capsule networks. 2018. URL: <http://papers.nips.cc/paper/8100-group-equivariant-capsule-networks.pdf>.
26. Phong N. H., Ribeiro B. Advanced capsule networks via context awareness. 2019. URL: <https://arxiv.org/pdf/1903.07497.pdf>.
27. Phay S., Sikka A., Dhall A., Bathula D. Dense and diverse capsule networks: Making the capsules learn better. 2018. URL: <https://arxiv.org/pdf/1805.04001.pdf>.
28. Neill J. O. Siamese capsule networks. 2018. URL: <https://arxiv.org/pdf/1805.07242.pdf>.
29. Liu Y., Zhang Q., Zhang D., Han J. Employing deep part-object relationships for salient object detection. 2019. URL: [https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Liu\\_Employing\\_Deep\\_Part-Object\\_Relationships\\_for\\_Salient\\_Object\\_Detection\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Liu_Employing_Deep_Part-Object_Relationships_for_Salient_Object_Detection_ICCV_2019_paper.pdf).

30. LaLonde R., Bagci U. Official Implementation of the Paper "Capsules for Object Segmentation". 2020. URL: <https://github.com/lalonderodney/SegCaps>.
31. Zhang W., Tang P., Zhao L. Remote sensing image scene classification using CNN-CapsNet. 2019. URL: <https://www.mdpi.com/2072-4292/11/5/494>.
32. Yang S., Lee F., Miao R., Cai J., Chen L., Yao W., Kotani K., Chen Q. RS-CapsNet: an advanced capsule network. 2020. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9086631>.
33. Wang L., Nie R., Xin R., Zhang J., Caia J. ScCapsNet: a deep learning classifier with the capability of interpretable feature extraction, applicable for single cell RNA data analysis. 2019. URL: <https://www.biorxiv.org/content/biorxiv/early/2019/05/20/506642.full.pdf>.
34. Nguyen H. H., Yamagishi J., Echizen I. Capsule-forensics: Using capsule networks to detect forged images and videos. 2019. URL: <https://arxiv.org/pdf/1810.11215.pdf>.
35. Katebi R., Zhou Y., Chornock R., Bunesco R. Galaxy morphology prediction using capsule networks. 2018. URL: <https://arxiv.org/pdf/1809.08377.pdf>.
36. Chattopadhyay A., Nabizadeh E., Hassanzadeh P. Analog Forecasting of Extreme-Causing Weather Patterns Using Deep Learning. 2020. URL: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001958>.
37. Al-Janabi S., Alkaim A. F., Adel Z. An Innovative synthesis of deep learning techniques (DCapsNet & DCOM) for generation electrical renewable energy from wind energy. 2020. URL: <https://link.springer.com/article/10.1007%2Fs00500-020-04905-9>.
38. Peng P., He Z., Wang L., Jiang Y. Microseismic records classification using capsule network with limited training samples in underground mining. 2020. URL: <https://www.nature.com/articles/s41598-020-70916-z>.
39. Jesus D. R., Cuevas J., Rivera W., Crivelli S. Capsule networks for protein structure classification and prediction. 2018. URL: <https://arxiv.org/pdf/1808.07475.pdf>.
40. Li Y., Qian M., Liu P., Cai Q., Xiaoying Li, Guo J., Yan H., Yu F., Yuan K., Yu J., Qin L., Liu H., Wu W., Xiao P., Zhou Z. The recognition of rice images by UAV based on capsule network. 2018. URL: <https://link.springer.com/article/10.1007%2Fs10586-018-2482-7>.
41. Altan G. Performance Evaluation of Capsule Networks for Classification of Plant Leaf Diseases. 2020. URL: <https://dergipark.org.tr/en/download/article-file/1299823>.
42. Sundarama D. M., Loganathan A. FSSCaps-DetCountNet: fuzzy soft sets and CapsNet-based detection and counting network for monitoring animals from aerial images. 2020. URL: <https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing-volume-14-issue-2/026521/FSSCaps-DetCountNet--fuzzy-soft-sets-and-CapsNet-based-detection/10.1117/1.JRS.14.026521.short?SSO=1>.
43. Kumar A. D., Karthika R., Parameswaran L. Novel deep learning model for traffic sign detection using capsule networks. – 2018. URL: <https://arxiv.org/ftp/arxiv/papers/1805/1805.04424.pdf>.
44. Kim Y., Wang P., Zhu Y., Mihaylova L. A capsule network for traffic speed prediction in complex road networks. 2018. URL: <https://arxiv.org/pdf/1807.10603.pdf>.
45. Andersen P. Deep reinforcement learning using capsules in advanced game environments. 2018. URL: <https://arxiv.org/pdf/1801.09597.pdf>.
46. Nguyen D. Q., Vu T., Nguyen T. D., Nguyen D. Q., Phung D. A capsule network-based embedding model for knowledge graph completion and search personalization. 2019. URL: <https://www.aclweb.org/anthology/N19-1226.pdf>.
47. Zhang N., Deng S., Sun Z., Chen X., Zhang W., Chen H. Attention-based capsule networks with dynamic routing for relation extraction. 2018. URL: <https://arxiv.org/pdf/1812.11321.pdf>.
48. Riedel S., Yao L., McCallum A. Modeling relations and their mentions without labeled text. – 2010. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.414.2202&rep=rep1&type=pdf>.
49. Sorokin D., Gurevych I. Contextaware representations for knowledge base relation extraction. 2017. URL: <https://www.aclweb.org/anthology/D17-1188.pdf>.
50. Beşer F., Kizrak M. A., Bolat B., Yildirim T. Recognition of sign language using capsule networks. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8404385>.
51. Suri K., Gupta R. Continuous sign language recognition from wearable IMUs using deep capsule networks and game theory. 2020. URL: <https://arxiv.org/ftp/arxiv/papers/2005/2005.00409.pdf>.
52. Rathnayaka P., Abeysinghe S., Samarajeewa C., Manchanayake I., Walpola M. Sentylic at IEST 2018: Gated recurrent neural network and capsule network based approach for implicit emotion detection. 2018. URL: <https://arxiv.org/pdf/1809.01452.pdf>.
53. Guo J., Fang F., Wang W., Ren F. EEG Emotion Recognition Based on Granger Causality and CapsNet Neural Network. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8691230>.
54. Chao H., Dong L., Liu Y., Lu B. Emotion recognition from multiband EEG signals using CapsNet. 2019. URL: <https://www.mdpi.com/1424-8220/19/9/2212/htm>.
55. Tereikovska L., Tereikovskiy I. Recognition of emotions by facial Geometry using a capsule neural network. 2019. URL: [https://d1wqtxts1xzl7.cloudfront.net/59388634/IJCIET\\_10\\_03\\_13920190525-126815-1bql6w1.pdf?1558777466=&response-content-disposition=inline%3B+filename%3DRECOGNITION\\_OF\\_EMOTIONS\\_BY\\_FACIAL\\_GEOMET.pdf&Expires=1602081766&Signature=FwBMBpFUGiebRjL-ZUswbT4maTfv6\\_HR2QTjn-3GBTBp\\_FnECJHEJw0sVlePFvt3p-wUGvliCXbSyIbVKF4oGMEI9DzQuNnhbFTGdaU-1jKP888hbirHQRc~xFFliGJx-fkEtnPn-Hq5kjpN3GSW90WQr9L8KES6SZfNO~wkorNpfW67pQJU1-fGDITkMm9Gp10sjkNKvGMrGXar9npefE7q-F3T9ETxyz66rkQH~HCHtdxkQjo0tHij75bnPRqgXCuFI7743um4eH-REcqioJ1bxWj2RNa4sJBBUJlhWjok4CSiqorZL~GV2fDgh9EmxGAL-nP70KP5oAO6Fwlmg\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzl7.cloudfront.net/59388634/IJCIET_10_03_13920190525-126815-1bql6w1.pdf?1558777466=&response-content-disposition=inline%3B+filename%3DRECOGNITION_OF_EMOTIONS_BY_FACIAL_GEOMET.pdf&Expires=1602081766&Signature=FwBMBpFUGiebRjL-ZUswbT4maTfv6_HR2QTjn-3GBTBp_FnECJHEJw0sVlePFvt3p-wUGvliCXbSyIbVKF4oGMEI9DzQuNnhbFTGdaU-1jKP888hbirHQRc~xFFliGJx-fkEtnPn-Hq5kjpN3GSW90WQr9L8KES6SZfNO~wkorNpfW67pQJU1-fGDITkMm9Gp10sjkNKvGMrGXar9npefE7q-F3T9ETxyz66rkQH~HCHtdxkQjo0tHij75bnPRqgXCuFI7743um4eH-REcqioJ1bxWj2RNa4sJBBUJlhWjok4CSiqorZL~GV2fDgh9EmxGAL-nP70KP5oAO6Fwlmg_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA).
56. Wu X., Liu S., Cao Y., Li X., Yu J., Dai D., Ma X., Hu S., Wu Z., Liu X., Meng H. Speech emotion recognition using capsule networks. 2019. URL: [https://www.researchgate.net/profile/Xixin\\_Wu/publication/n331298406\\_Speech\\_Emotion\\_Recognition\\_Using\\_Capsule\\_Networks/links/5d0a41ac458515ea1a70f2b1/Speech-Emotion-Recognition-Using-Capsule-Networks.pdf](https://www.researchgate.net/profile/Xixin_Wu/publication/n331298406_Speech_Emotion_Recognition_Using_Capsule_Networks/links/5d0a41ac458515ea1a70f2b1/Speech-Emotion-Recognition-Using-Capsule-Networks.pdf).
57. Xia C., Zhang C., Yan X., Chang Y., Yu P. S. Zero-shot user intent detection via capsule neural networks. 2018. URL: <https://arxiv.org/pdf/1809.00385.pdf>.



58. Li C., Quan C., Peng L., Qi Y., Deng Y., Wu L. A Capsule Network for recommendation and explaining what you like and dislike. 2019. URL: <https://arxiv.org/pdf/1907.00687.pdf>.
59. Duarte K., Rawat Y., Shah M. VideoCapsuleNet: A simplified network for action detection. 2018. URL: <http://papers.nips.cc/paper/7988-videoCapsuleNet-a-simplified-network-for-action-detection.pdf>.
60. McIntosh B., Duarte K., Rawat Y. S., Shah M. Visual-Textual Capsule Routing for Text-Based Video Segmentation. 2020. URL: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/papers/McIntosh\\_Visual-Textual\\_Capsule\\_Routing\\_for\\_Text-Based\\_Video\\_Segmentation\\_CVPR\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2020/papers/McIntosh_Visual-Textual_Capsule_Routing_for_Text-Based_Video_Segmentation_CVPR_2020_paper.pdf).
61. Goldani M. H., Momtazi S., Safabakhsh R. Detecting Fake News with Capsule Neural Networks. 2020. URL: <https://arxiv.org/pdf/2002.01030.pdf>.
62. Bonan L., Tong Z., Min W. Fast exact classification algorithm of massive fingerprint patterns based on capsule network. 2020. URL: <http://html.rhhz.net/ZGKXYDXXB/20200312.htm>.
63. Ning R. Secure Mobile Computing by Using Convolutional and Capsule Deep Neural Networks. – 2020. URL: [https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1219&context=ece\\_etds](https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1219&context=ece_etds).
64. Iqbal T., Xu Y., Kong Q., Wang W. Capsule routing for sound event detection. 2018. URL: <https://arxiv.org/pdf/1806.04699.pdf>.
65. Jain R. Improving performance and inference on audio classification tasks using capsule networks. 2019. URL: <https://arxiv.org/pdf/1902.05069.pdf>.
66. Bonheur S., Stern D., Payer C., Pienn M., Olschewski H., Urschler M. Matwo-CapsNet: A Multi-label semantic segmentation capsules network. 2019. URL: <https://graz.pure.elsevier.com/en/publications/matwo-capsnet-a-multi-label-semantic-segmentation-capsules-network>.
67. Jayasundara V., Jayasekara S., Jayasekara H., Rajasegaran J., Seneviratne S., Rodrigo R. Textcaps: Handwritten character recognition with very small datasets. 2019. URL: <https://arxiv.org/pdf/1904.08095.pdf>.
68. Zhao Y., Bidal T., Deng H., Tombari F. 3D Point capsule networks. 2019. URL: [https://openaccess.thecvf.com/content\\_CVPR\\_2019/papers/Zhao\\_3D\\_Point\\_Capsule\\_Networks\\_CVPR\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2019/papers/Zhao_3D_Point_Capsule_Networks_CVPR_2019_paper.pdf).
69. Paoletti M. E., Haut J. M., Fernandez-Beltran R. Capsule networks for hyperspectral image classification. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8509610>.
70. Jimenez-Sanchez A., Albarqouni S., Mateus D. Capsule networks against medical imaging data challenges. 2018. URL: <https://arxiv.org/pdf/1807.07559.pdf>.
71. Kruthika K. R., Maheshappa H. D. CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis. 2019. URL: <https://www.sciencedirect.com/science/article/pii/S235291481830176X>.
72. Afshar P., Mohammadi A., Plataniotis K. N. Brain tumor type classification via capsule networks. 2018. URL: <https://arxiv.org/pdf/1802.10200.pdf>.
73. Mobiny A., Lu H., Nguyen H. V., Roysam B., Varadarajan N. Automated classification of apoptosis in phase contrast microscopy using capsule network. 2019. URL: <https://par.nsf.gov/servlets/purl/10110677>.
74. Mobiny A., Nguyen H. Fast CapsNet for lung cancer screening. 2018. URL: <https://arxiv.org/pdf/1806.07416.pdf>.
75. Mitta A., Kumar D., Mittal M., Saba T., Abunadi I., Rehman A., Roy S. Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images. 2020. URL: <https://www.mdpi.com/1424-8220/20/4/1068/htm>.
76. Sandu N., Karim S. The application of fast CapsNet computer vision in detecting Covid-19. 2020. URL: [https://www.researchgate.net/profile/Raj\\_Sandu2/publication/342699125\\_The\\_application\\_of\\_fast\\_CapsNet\\_computer\\_vision\\_in\\_detecting\\_Covid-19/links/5f01b753299bf18816037ea0/The-application-of-fast-CapsNet-computer-vision-in-detecting-Covid-19.pdf](https://www.researchgate.net/profile/Raj_Sandu2/publication/342699125_The_application_of_fast_CapsNet_computer_vision_in_detecting_Covid-19/links/5f01b753299bf18816037ea0/The-application-of-fast-CapsNet-computer-vision-in-detecting-Covid-19.pdf).
77. Iesmantas T., Alzbutas R. Convolutional capsule network for classification of breast cancer histology images. 2018. URL: <https://arxiv.org/ftp/arxiv/papers/1804/1804.08376.pdf>.
78. Maity M., Jaiswal A., Gantait K., Chatterjee J., Mukherjee A., Quantification of malaria parasitaemia using trainable semantic segmentation and capsnet. 2020. URL: [https://www.researchgate.net/profile/Kripasindhu\\_Gantait/publication/342770757\\_Quantification\\_of\\_malaria\\_parasitaemia\\_using\\_trainable\\_semantic\\_segmentation\\_and\\_capsnet/links/5f0effde299b1e548b71167/Quantification-of-malaria-parasitaemia-using-trainable-semantic-segmentation-and-capsnet.pdf](https://www.researchgate.net/profile/Kripasindhu_Gantait/publication/342770757_Quantification_of_malaria_parasitaemia_using_trainable_semantic_segmentation_and_capsnet/links/5f0effde299b1e548b71167/Quantification-of-malaria-parasitaemia-using-trainable-semantic-segmentation-and-capsnet.pdf).
79. Kose U., Deperlioglu O., Alzubi J., Patrut B. Diagnosing of Diabetic Retinopathy with Image Dehazing and Capsule Network. 2020. URL: [https://link.springer.com/chapter/10.1007/978-981-15-6325-6\\_9](https://link.springer.com/chapter/10.1007/978-981-15-6325-6_9).
80. Butun E., Yildirim O., Talo M., Tanc R., Acharya R. 1D-CADCapsNet: One dimensional deep capsule networks for coronary artery disease detection using ECG signals. 2020. URL: [https://www.researchgate.net/profile/Oezal\\_Yildirim/publication/338452030\\_1D-CADCapsNet\\_One\\_Dimensional\\_Deep\\_Capsule\\_Networks\\_for\\_Coronary\\_Artery\\_Disease\\_Detection\\_Using\\_ECG\\_Signals/links/5e15c7dda6fdcc283761fac8/1D-CADCapsNet-One-Dimensional-Deep-Capsule-Networks-for-Coronary-Artery-Disease-Detection-Using-ECG-Signals.pdf](https://www.researchgate.net/profile/Oezal_Yildirim/publication/338452030_1D-CADCapsNet_One_Dimensional_Deep_Capsule_Networks_for_Coronary_Artery_Disease_Detection_Using_ECG_Signals/links/5e15c7dda6fdcc283761fac8/1D-CADCapsNet-One-Dimensional-Deep-Capsule-Networks-for-Coronary-Artery-Disease-Detection-Using-ECG-Signals.pdf).
81. Janeera D. A., Sasipriya S. A Brain Computer Interface Based Patient Observation and Indoor Locating System with Capsule Network Algorithm. 2020. URL: [https://link.springer.com/chapter/10.1007/978-3-030-51859-2\\_23](https://link.springer.com/chapter/10.1007/978-3-030-51859-2_23).
82. Lin A., Li J., Ma Z. On Learning and Learned Representation with Dynamic Routing in Capsule Networks. 2020. URL: <https://arxiv.org/pdf/1810.04041.pdf>.
83. Ohta N., Kawai S., Nobuhara H. Analysis and Learning of Capsule Networks Robust for Small Image Deformation. 2020. URL: <https://ieeexplore.ieee.org/abstract/document/9206651>.
84. Ren H., Su J., Lu H. Evaluating Generalization Ability of Convolutional Neural Networks and Capsule Networks for Image Classification via Top-2 Classification. 2019. URL: <https://arxiv.org/pdf/1901.10112.pdf>.
85. Chauhan A., Babu M., Kandru N., Lokegaonkar S. Empirical Study on convergence of Capsule Networks with various hyperparameters. 2018. URL: <http://people.cs.vt.edu/~bhuang/courses/opt18/projects/capsule.pdf>.
86. Jacob B., Kligys S., Chen B., Zhu M., Tang M., Howard A., Adam H., Kalenichenko D. Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. 2017. URL: <https://arxiv.org/pdf/1712.05877.pdf>.

## References

1. Sabour S., Frosst N., G. Hinton. Dynamic routing between capsules. 2017. URL: <http://papers.nips.cc/paper/6975-dynamic-routing-between-capsules.pdf>.
2. Shahroudjeh A., Mohammadi A., Plataniotis K. N. Improved explainability of capsule networks: relevance path by agreement. 2018. URL: <https://arxiv.org/pdf/1802.10204.pdf>.
3. Hinton G., Sabou S., Frosst N. Matrix capsules with EM routing. 2018. URL: <https://openreview.net/pdf?id=HJWLfGWRb>.
4. Zhang S., Zhao W., Wu X., Zhou Q. Fast dynamic routing based on weighted kernel density estimation. 2018. URL: <https://arxiv.org/pdf/1805.10807.pdf>.
5. Choi J., Seo H., Im S., Kang M. Attention routing between capsules. 2019. URL: [https://openaccess.thecvf.com/content\\_ICCVW\\_2019/papers/NeurArch/Choi\\_Attention\\_Routing\\_Between\\_Capsules\\_ICCVW\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCVW_2019/papers/NeurArch/Choi_Attention_Routing_Between_Capsules_ICCVW_2019_paper.pdf).
6. Huang W., Zhou F. DA-CapsNet: dual attention mechanism capsule network. 2020. URL: <https://www.nature.com/articles/s41598-020-68453-w>.
7. Li H., Guo X., Dai B., Ouyang W., Wang X. Neural network encapsulation. 2018. URL: <https://arxiv.org/pdf/1808.03749.pdf>.
8. Li H., Wang Y. Cognitive consistency routing algorithm of capsule-network. 2018. URL: <https://arxiv.org/pdf/1808.09062.pdf>.
9. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015. URL: <https://arxiv.org/pdf/1502.03167.pdf>.
10. Chen Z., Crandall D. Generalized capsule networks with trainable routing procedure. 2018. URL: <https://arxiv.org/pdf/1808.08692.pdf>.
11. Peer D., Stabinger S., Rodriguez-Sanchez A. Training deep capsule networks. 2018. URL: [https://www.researchgate.net/profile/Sebastian\\_Stabinger/publication/329945773\\_Training\\_Deep\\_Capsule\\_Networks/links/5c7e8901299b1268d3a8928/Training-Deep-Capsule-Networks.pdf](https://www.researchgate.net/profile/Sebastian_Stabinger/publication/329945773_Training_Deep_Capsule_Networks/links/5c7e8901299b1268d3a8928/Training-Deep-Capsule-Networks.pdf).
12. Jia B., Huang Q. DE-CapsNet: A diverse enhanced capsule network with disperse dynamic. 2020. URL: <https://www.mdpi.com/2076-3417/10/3/884>.
13. Zhao Z., Kleinhans A., Sandhu G., Patel I., Unnikrishnanra K. P. Capsule networks with max-min normalization. 2019. URL: <https://arxiv.org/pdf/1903.09662.pdf>.
14. Wang D., Liu Q. An optimization view on dynamic routing between capsules. 2018. URL: <https://openreview.net/pdf?id=HJtFYJdf>.
15. Ramasinghe S., Athuraliya C. D., Khan S. H. A Context-aware Capsule Network for Multi-label Classification. 2018. URL: [https://openaccess.thecvf.com/content\\_ECCVW\\_2018/papers/11131/Ramasinghe\\_A\\_Context-aware\\_Capsule\\_Network\\_for\\_Multi-label\\_Classification\\_ECCVW\\_2018\\_paper.pdf](https://openaccess.thecvf.com/content_ECCVW_2018/papers/11131/Ramasinghe_A_Context-aware_Capsule_Network_for_Multi-label_Classification_ECCVW_2018_paper.pdf).
16. Ribeiro F., Leontidis G., Kollias S. Capsule routing via variational Bayes. 2019. URL: <https://arxiv.org/pdf/1905.11455.pdf>.
17. LaLonde R., Bagci U. Capsules for object segmentation. 2018. URL: <https://arxiv.org/abs/1804.04241>.
18. Chen Z., Li X., Wang C., Crandall D. Capsule networks without routing procedures. 2019. URL: <https://openreview.net/pdf?id=BlgNfkrYvS>.
19. Bahadori M. T. Spectral capsule networks. 2018. URL: <https://openreview.net/pdf?id=HJuMvYPaM>.
20. Zhang L., Edraki M., Qi G. J. CapProNet: Deep feature learning via orthogonal projections onto capsule subspaces. 2018. URL: <http://papers.nips.cc/paper/7823-capronet-deep-feature-learning-via-orthogonal-projections-onto-capsule-subspaces.pdf>.
21. Xiang C., Zhang L., Tang Y., Zou W., Xu C. MS-CapsNet: A novel multi-scale capsule network. 2018. URL: [https://hal.archives-ouvertes.fr/hal-01908269/file/Xiang%20et%20al\\_2018\\_MS-CapsNet.pdf](https://hal.archives-ouvertes.fr/hal-01908269/file/Xiang%20et%20al_2018_MS-CapsNet.pdf).
22. Jaiswal A., AbdAlmageed W., Wu Y., Natarajan P. CapsuleGAN: Generative adversarial capsule network. 2018. URL: [https://openaccess.thecvf.com/content\\_ECCVW\\_2018/papers/11131/Jaiswal\\_CapsuleGAN\\_Generative\\_Adversarial\\_Capsule\\_Network\\_ECCVW\\_2018\\_paper.pdf](https://openaccess.thecvf.com/content_ECCVW_2018/papers/11131/Jaiswal_CapsuleGAN_Generative_Adversarial_Capsule_Network_ECCVW_2018_paper.pdf).
23. Saqr R., Vivona S. CapsGAN: Using dynamic routing for generative adversarial networks. 2018. URL: <https://arxiv.org/pdf/1806.03968.pdf>.
24. Majdabadi M. M., Ko S. B. MSG-CapsGAN: Multi-scale gradient capsule GAN for face super resolution. 2020. URL: <https://ieeexplore.ieee.org/abstract/document/9051244>.
25. Lenssen J. E., Fey M., Libuschewski P. Group equivariant capsule networks. 2018. URL: <http://papers.nips.cc/paper/8100-group-equivariant-capsule-networks.pdf>.
26. Phong N. H., Ribeiro B. Advanced capsule networks via context awareness. 2019. URL: <https://arxiv.org/pdf/1903.07497.pdf>.
27. Phaye S., Sikka A., Dhall A., Bathula D. Dense and diverse capsule networks: Making the capsules learn better. 2018. URL: <https://arxiv.org/pdf/1805.04001.pdf>.
28. Neill J. O. Siamese capsule networks. 2018. URL: <https://arxiv.org/pdf/1805.07242.pdf>.
29. Liu Y., Zhang Q., Zhang D., Han J. Employing deep part-object relationships for salient object detection. 2019. URL: [https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Liu\\_Employing\\_Deep\\_Part-Object\\_Relationships\\_for\\_Salient\\_Object\\_Detection\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Liu_Employing_Deep_Part-Object_Relationships_for_Salient_Object_Detection_ICCV_2019_paper.pdf).
30. LaLonde R., Bagci U. Official Implementation of the Paper "Capsules for Object Segmentation". 2020. URL: <https://github.com/lalonderodney/SegCaps>.
31. Zhang W., Tang P., Zhao L. Remote sensing image scene classification using CNN-CapsNet. 2019. URL: <https://www.mdpi.com/2072-4292/11/5/494>.
32. Yang S., Lee F., Miao R., Cai J., Chen L., Yao W., Kotani K., Chen Q. RS-CapsNet: an advanced capsule network. 2020. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9086631>.
33. Wang L., Nie R., Xin R., Zhange J., Caia J. ScCapsNet: a deep learning classifier with the capability of interpretable feature extraction, applicable for single cell RNA data analysis. 2019. URL: <https://www.biorxiv.org/content/biorxiv/early/2019/05/20/506642.full.pdf>.
34. Nguyen H. H., Yamagishi J., Echizen I. Capsule-forensics: Using capsule networks to detect forged images and videos. 2019. URL: <https://arxiv.org/pdf/1810.11215.pdf>.
35. Katebi R., Zhou Y., Chornock R., Bunesu R. Galaxy morphology prediction using capsule networks. 2018. URL: <https://arxiv.org/pdf/1809.08377.pdf>.
36. Chattopadhyay A., Nabizadeh E., Hassanzadeh P. Analog Forecasting of Extreme-Causing Weather Patterns Using Deep Learning. 2020. URL: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001958>.

37. Al-Janabi S., Alkaim A. F., Adel Z. An Innovative synthesis of deep learning techniques (DCapsNet & DCOM) for generation electrical renewable energy from wind energy. 2020. URL: <https://link.springer.com/article/10.1007%2Fs00500-020-04905-9>.
38. Peng P., He Z., Wang L., Jiang Y. Microseismic records classification using capsule network with limited training samples in underground mining. 2020. URL: <https://www.nature.com/articles/s41598-020-70916-z>
39. Jesus D. R., Cuevas J., Rivera W., Crivelli S. Capsule networks for protein structure classification and prediction. 2018. URL: <https://arxiv.org/pdf/1808.07475.pdf>.
40. Li Y., Qian M., Liu P., Cai Q., Xiaoying Li, Guo J., Yan H., Yu F., Yuan K., Yu J., Qin L., Liu H., Wu W., Xiao P., Zhou Z. The recognition of rice images by UAV based on capsule network. 2018. URL: <https://link.springer.com/article/10.1007%2Fs10586-018-2482-7>.
41. Altan G. Performance Evaluation of Capsule Networks for Classification of Plant Leaf Diseases. 2020. URL: <https://dergipark.org.tr/en/download/article-file/1299823>.
42. Sundarama D. M., Loganathanb A. FSSCaps-DetCountNet: fuzzy soft sets and CapsNet-based detection and counting network for monitoring animals from aerial images. 2020. URL: <https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing/volume-14/issue-2/026521/FSSCaps-DetCountNet--fuzzy-soft-sets-and-CapsNet-based-detection/10.1117/1.JRS.14.026521.short?SSO=1>.
43. Kumar A. D., Karthika R., Parameswaran L. Novel deep learning model for traffic sign detection using capsule networks. – 2018. URL: <https://arxiv.org/ftp/arxiv/papers/1805/1805.04424.pdf>.
44. Kim Y., Wang P., Zhu Y., Mihaylova L. A capsule network for traffic speed prediction in complex road networks. 2018. URL: <https://arxiv.org/pdf/1807.10603.pdf>.
45. Andersen P. Deep reinforcement learning using capsules in advanced game environments. 2018. URL: <https://arxiv.org/pdf/1801.09597.pdf>.
46. Nguyen D. Q., Vu T., Nguyen T. D., Nguyen D. Q., Phung D. A capsule network-based embedding model for knowledge graph completion and search personalization. 2019. URL: <https://www.aclweb.org/anthology/N19-1226.pdf>.
47. Zhang N., Deng S., Sun Z., Chen X., Zhang W., Chen H. Attention-based capsule networks with dynamic routing for relation extraction. 2018. URL: <https://arxiv.org/pdf/1812.11321.pdf>.
48. Riedel S., Yao L., McCallum A. Modeling relations and their mentions without labeled text. – 2010. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.414.2202&rep=rep1&type=pdf>.
49. Sorokin D., Gurevych I. Contextaware representations for knowledge base relation extraction. 2017. URL: <https://www.aclweb.org/anthology/D17-1188.pdf>.
50. Beşer F., Kizrak M. A., Bolat B., Yildirim T. Recognition of sign language using capsule networks. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8404385>.
51. Suri K., Gupta R. Continuous sign language recognition from wearable IMUs using deep capsule networks and game theory. 2020. URL: <https://arxiv.org/ftp/arxiv/papers/2005/2005.00409.pdf>.
52. Rathnayaka P., Abeysinghe S., Samarajeewa C., Manchanayake I., Walpola M. Sentylic at IEST 2018: Gated recurrent neural network and capsule network based approach for implicit emotion detection. 2018. URL: <https://arxiv.org/pdf/1809.01452.pdf>.
53. Guo J., Fang F., Wang W., Ren F. EEG Emotion Recognition Based on Granger Causality and CapsNet Neural Network. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8691230>.
54. Chao H., Dong L., Liu Y., Lu B. Emotion recognition from multiband EEG signals using CapsNet. 2019. URL: <https://www.mdpi.com/1424-8220/19/9/2212/htm>.
55. Tereikovska L., Tereikovskiy I. Recognition of emotions by facial Geometry using a capsule neural network. 2019. URL: [https://d1wqtxts1xzle7.cloudfront.net/59388634/IJCIET\\_10\\_03\\_13920190525-126815-1bql6w1.pdf?1558777466=&response-content-disposition=inline%3B+filename%3DRECOGNITION\\_OF\\_EMOTIONS\\_BY\\_FACIAL\\_GEOMET.pdf&Expires=1602081766&Signature=FWMbpfUGiebRjl-ZUswbT4maTfv6HR2QTjn-3GBTBpFnECJHEJw0sVlePFvt3p-wUGvliCXbSyIbVKF4oGMEIg9DzQuNnhbxFtGdaU-ljKP888hbirHQRC~xFFliGJxF-kEtfnPn-Hq5kjpN3GSW90WQr9L8KES6SZfNO~wkorNpFW67pQJUI~fGDITkMm9Gp10sjkNKvGMrGXar9npefE7q-F3T9ETxyz66rkQH~HCHtdxkQjo0tHij75bnPRqgXCuFI7743um4eH-REcqioJ1bxWj2RNa4sJBBUJlhWjok4CSiqorZL~GV2fDg h9EmxGAL-nP70KP5oAO6Fwlmg\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/59388634/IJCIET_10_03_13920190525-126815-1bql6w1.pdf?1558777466=&response-content-disposition=inline%3B+filename%3DRECOGNITION_OF_EMOTIONS_BY_FACIAL_GEOMET.pdf&Expires=1602081766&Signature=FWMbpfUGiebRjl-ZUswbT4maTfv6HR2QTjn-3GBTBpFnECJHEJw0sVlePFvt3p-wUGvliCXbSyIbVKF4oGMEIg9DzQuNnhbxFtGdaU-ljKP888hbirHQRC~xFFliGJxF-kEtfnPn-Hq5kjpN3GSW90WQr9L8KES6SZfNO~wkorNpFW67pQJUI~fGDITkMm9Gp10sjkNKvGMrGXar9npefE7q-F3T9ETxyz66rkQH~HCHtdxkQjo0tHij75bnPRqgXCuFI7743um4eH-REcqioJ1bxWj2RNa4sJBBUJlhWjok4CSiqorZL~GV2fDg h9EmxGAL-nP70KP5oAO6Fwlmg_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA).
56. Wu X., Liu S., Cao Y., Li X., Yu J., Dai D., Ma X., Hu S., Wu Z., Liu X., Meng H. Speech emotion recognition using capsule networks. 2019. URL: [https://www.researchgate.net/profile/Xixin\\_Wu/publication/n331298406\\_Speech\\_Emotion\\_Recognition\\_Using\\_Capsule\\_Networks/links/5d0a41ac458515ea1a70f2b1/Speech-Emotion-Recognition-Using-Capsule-Networks.pdf](https://www.researchgate.net/profile/Xixin_Wu/publication/n331298406_Speech_Emotion_Recognition_Using_Capsule_Networks/links/5d0a41ac458515ea1a70f2b1/Speech-Emotion-Recognition-Using-Capsule-Networks.pdf).
57. Xia C., Zhang C., Yan X., Chang Y., Yu P. S. Zero-shot user intent detection via capsule neural networks. 2018. URL: <https://arxiv.org/pdf/1809.00385.pdf>.
58. Li C., Quan C., Peng L., Qi Y., Deng Y., Wu L. A Capsule Network for recommendation and explaining what you like and dislike. 2019. URL: <https://arxiv.org/pdf/1907.00687.pdf>.
59. Duarte K., Rawat Y., Shah M. VideoCapsuleNet: A simplified network for action detection. 2018. URL: <http://papers.nips.cc/paper/7988-videocapsulenet-a-simplified-network-for-action-detection.pdf>.
60. McIntosh B., Duarte K., Rawat Y. S., Shah M. Visual-Textual Capsule Routing for Text-Based Video Segmentation. 2020. URL: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/papers/McIntosh\\_Visual-Textual\\_Capsule\\_Routing\\_for\\_Text-Based\\_Video\\_Segmentation\\_CVPR\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2020/papers/McIntosh_Visual-Textual_Capsule_Routing_for_Text-Based_Video_Segmentation_CVPR_2020_paper.pdf).
61. Goldani M. H., Momtazi S., Safabakhsh R. Detecting Fake News with Capsule Neural Networks. 2020. URL: <https://arxiv.org/pdf/2002.01030.pdf>.
62. Bonan L., Tong Z., Min W. Fast exact classification algorithm of massive fingerprint patterns based on capsule network. 2020. URL: <http://html.rhhz.net/ZGKXYDXXB/20200312.htm>.
63. Ning R. Secure Mobile Computing by Using Convolutional and Capsule Deep Neural Networks. – 2020. URL: [https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1219&context=ece\\_etds](https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1219&context=ece_etds).

64. Iqbal T., Xu Y., Kong Q., Wang W. Capsule routing for sound event detection. 2018. URL: <https://arxiv.org/pdf/1806.04699.pdf>.
65. Jain R. Improving performance and inference on audio classification tasks using capsule networks. 2019. URL: <https://arxiv.org/pdf/1902.05069.pdf>.
66. Bonheur S., Stern D., Payer C., Pienn M., Olschewski H., Urschler M. Matwo-CapsNet: A Multi-label semantic segmentation capsules network. 2019. URL: <https://graz.pure.elsevier.com/en/publications/matwo-capsnet-a-multi-label-semantic-segmentation-capsules-network>.
67. Jayasundara V., Jayasekara S., Jayasekara H., Rajasegaran J., Seneviratne S., Rodrigo R. Textcaps: Handwritten character recognition with very small datasets. 2019. URL: <https://arxiv.org/pdf/1904.08095.pdf>.
68. Zhao Y., Birdal T., Deng H., Tombari F. 3D Point capsule networks. 2019. URL: [https://openaccess.thecvf.com/content\\_CVPR\\_2019/papers/Zhao\\_3D\\_Point\\_Capsule\\_Networks\\_CVPR\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2019/papers/Zhao_3D_Point_Capsule_Networks_CVPR_2019_paper.pdf).
69. Paoletti M. E., Haut J. M., Fernandez-Beltran R. Capsule networks for hyperspectral image classification. 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8509610>.
70. Jimenez-Sanchez A., Albarqouni S., Mateus D. Capsule networks against medical imaging data challenges. 2018. URL: <https://arxiv.org/pdf/1807.07559.pdf>.
71. Kruthika K. R., Maheshappa H. D. CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis. 2019. URL: <https://www.sciencedirect.com/science/article/pii/S235291481830176X>.
72. Afshar P., Mohammadi A., Plataniotis K. N. Brain tumor type classification via capsule networks. 2018. URL: <https://arxiv.org/pdf/1802.10200.pdf>.
73. Mobiny A., Lu H., Nguyen H. V., Roysam B., Varadarajan N. Automated classification of apoptosis in phase contrast microscopy using capsule network. 2019. URL: <https://par.nsf.gov/servlets/purl/10110677>.
74. Mobiny A., Nguyen H. Fast CapsNet for lung cancer screening. 2018. URL: <https://arxiv.org/pdf/1806.07416.pdf>.
75. Mitta A., Kumar D., Mittal M., Saba T., Abunadi I., Rehman A., Roy S. Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images. 2020. URL: <https://www.mdpi.com/1424-8220/20/4/1068/htm>.
76. Sandu N., Karim S. The application of fast CapsNet computer vision in detecting Covid-19. 2020. URL: [https://www.researchgate.net/profile/Raj\\_Sandu2/publication/342699125\\_The\\_application\\_of\\_fast\\_CapsNet\\_computer\\_vision\\_in\\_detecting\\_Covid-19/links/5f01b753299bf18816037ea0/The-application-of-fast-CapsNet-computer-vision-in-detecting-Covid-19.pdf](https://www.researchgate.net/profile/Raj_Sandu2/publication/342699125_The_application_of_fast_CapsNet_computer_vision_in_detecting_Covid-19/links/5f01b753299bf18816037ea0/The-application-of-fast-CapsNet-computer-vision-in-detecting-Covid-19.pdf).
77. Iesmantas T., Alzbutas R. Convolutional capsule network for classification of breast cancer histology images. 2018. URL: <https://arxiv.org/ftp/arxiv/papers/1804/1804.08376.pdf>.
78. Maity M., Jaiswal A., Gantait K., Chatterjee J., Mukherjee A., Quantification of malaria parasitaemia using trainable semantic segmentation and capsnet. 2020. URL: [https://www.researchgate.net/profile/Kripasindhu\\_Gantait/publication/342770757\\_Quantification\\_of\\_malaria\\_parasitaemia\\_using\\_trainable\\_semantic\\_segmentation\\_and\\_capsnet/links/5f0effde299bf1e548b71167/Quantification-of-malaria-parasitaemia-using-trainable-semantic-segmentation-and-capsnet.pdf](https://www.researchgate.net/profile/Kripasindhu_Gantait/publication/342770757_Quantification_of_malaria_parasitaemia_using_trainable_semantic_segmentation_and_capsnet/links/5f0effde299bf1e548b71167/Quantification-of-malaria-parasitaemia-using-trainable-semantic-segmentation-and-capsnet.pdf).
79. Kose U., Deperlioglu O., Alzubi J., Patrut B. Diagnosing of Diabetic Retinopathy with Image Dehazing and Capsule Network. 2020. URL: [https://link.springer.com/chapter/10.1007/978-981-15-6325-6\\_9](https://link.springer.com/chapter/10.1007/978-981-15-6325-6_9).
80. Butun E., Yildirim O., Talo M., Tanc R., Acharya R. 1D-CADCapsNet: One dimensional deep capsule networks for coronary artery disease detection using ECG signals. 2020. URL: [https://www.researchgate.net/profile/Oezal\\_Yildirim/publication/338452030\\_1D-CADCapsNet\\_One\\_Dimensional\\_Deep\\_Capsule\\_Networks\\_for\\_Coronary\\_Artery\\_Disease\\_Detection\\_Using\\_ECG\\_Signals/links/5e15c7dda6fdcc283761fac8/1D-CADCapsNet-One-Dimensional-Deep-Capsule-Networks-for-Coronary-Artery-Disease-Detection-Using-ECG-Signals.pdf](https://www.researchgate.net/profile/Oezal_Yildirim/publication/338452030_1D-CADCapsNet_One_Dimensional_Deep_Capsule_Networks_for_Coronary_Artery_Disease_Detection_Using_ECG_Signals/links/5e15c7dda6fdcc283761fac8/1D-CADCapsNet-One-Dimensional-Deep-Capsule-Networks-for-Coronary-Artery-Disease-Detection-Using-ECG-Signals.pdf).
81. Janeera D. A., Sasipriya S. A Brain Computer Interface Based Patient Observation and Indoor Locating System with Capsule Network Algorithm. 2020. URL: [https://link.springer.com/chapter/10.1007/978-3-030-51859-2\\_23](https://link.springer.com/chapter/10.1007/978-3-030-51859-2_23).
82. Lin A., Li J., Ma Z. On Learning and Learned Representation with Dynamic Routing in Capsule Networks. 2020. URL: <https://arxiv.org/pdf/1810.04041.pdf>.
83. Ohta N., Kawai S., Nobuhara H. Analysis and Learning of Capsule Networks Robust for Small Image Deformation. 2020. URL: <https://ieeexplore.ieee.org/abstract/document/9206651>.
84. Ren H., Su J., Lu H. Evaluating Generalization Ability of Convolutional Neural Networks and Capsule Networks for Image Classification via Top-2 Classification. 2019. URL: <https://arxiv.org/pdf/1901.10112.pdf>.
85. Chauhan A., Babu M., Kandru N., Lokegaonkar S. Empirical Study on convergence of Capsule Networks with various hyperparameters. 2018. URL: <http://people.cs.vt.edu/~bhuang/courses/opt18/projects/capsule.pdf>.
86. Jacob B., Kligys S., Chen B., Zhu M., Tang M., Howard A., Adam H., Kalenichenko D. Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. 2017. URL: <https://arxiv.org/pdf/1712.05877.pdf>.

**Арсенюк В.С., Барбарук В.Н. Анализ возможностей капсульных нейронных сетей для решения проблем, связанных с распознаванием изображений**

*В статье выполнен обзор публикаций о разработках модификаций базовых алгоритмов маршрутизации, базовой архитектуры капсульной нейросети и используемых данным сетями наборов данных. На основании анализа и обобщений сделаны выводы о принципах построения таких сетей на современном этапе, выборе алгоритма маршрутизации и оптимизатора для сложных датасетов.*

*Рассмотрен вопрос о возможности применения капсульного подхода при построении нейронной сети, распознающей изображения, для мобильных приложений.*

**Ключевые слова:** распознавание изображений, повышение точности определения объектов, капсульная нейронная сеть, мобильное приложение, облегченные модели.

**Arsenyuk V.S., Barbaruk V.M. Analysis of capsule neural network capabilities for solving image recognition problems.**

*When using convolutional neural networks as a classifier, the accuracy of the trained model may deteriorate in the case of illumination, viewing angle, or classification of real objects.*

*To solve this problem, CNN architectures become more complex by increasing the number of convolutional layers and feature maps that take into account scale, orientation, and so on. But striving to cover options and avoid overfitting leads to a significant increase in the amount of data for training and an increase in the size of the overtrained model. All this reduces the possibility of integrating such models into mobile applications. In addition, this model training approach tends to memorize data rather than generalize solutions.*

*The inability of convolutional neural networks to recognize posture, texture, rotation, deformation of the image as a whole or part of it, translational invariance, the presence of the «Picasso problem» – lead to the appearance of capsular neural networks.*

*Over the past two years, a significant amount of information has been published about the capsule approach to building neural networks. The article analyzes and summarizes research data, taking into account the possibility of using capsule networks as a tool to improve the accuracy of object recognition.*

*The authors reviewed publications on the development of modifications of the basic routing algorithms, the basic architecture of the capsule neural network and the datasets used by these networks, made conclusions about the principles of building such networks at the present stage, the choice of the routing algorithm and optimizer for complex datasets.*

*The question of the possibility of using the capsule approach when building a neural network that recognizes images for mobile applications is considered.*

*For overtrained mobile models, there are a number of requirements for the size of the models and the calculations that the model will perform. Size is limited as the model is loaded into RAM and consumes significant GPU and / or CPU computing resources. The computations that the model performs should be efficient enough to process large amounts of information in a reasonable amount of time, but without significant power consumption or heating up the device's battery.*

*This must be taken into account when creating neural network architecture, choosing routing algorithms. Why use the so-called «lightweight» models.*

*In addition, in order to reduce the size of the trained model, it is necessary to optimize the model for inference, to perform the operations of «freezing» and post-training quantization.*

**Keywords:** *image recognition, increasing the accuracy of object detection, capsule neural network, mobile application, lightweight models.*

**Арсенюк В.С.** – аспірант кафедри комп'ютерних наук та інженерії Східноукраїнського національного університету імені Володимира Даля, e-mail: vlad.arsenyuk@gmail.com

**Барбарук В.М.** – к.т.н., доцент, директор Центру удосконалення освіти Східноукраїнського національного університету імені Володимира Даля, e-mail: barbaruk.viktor@gmail.com

Стаття подана 05.11.2020.