



مجلة بحوث

جامعة حلب في المناطق المحررة

المجلد الثالث - العدد الثاني

الجزء الثاني

1445 / 12 / 7 هـ - 2024 / 6 / 13 م

علمية - ربعية - محكمة

تصدر عن

جامعة حلب في المناطق المحررة



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الهيئة الاستشارية لمجلة جامعة حلب في المناطق المحررة

د. جلال الدين خانجي أ.د. زكريا ظلام أ.د. عبد الكريم بكار
أ.د. إبراهيم أحمد الديبو أ.د. أسامة اختيار د. أسامة القاضي
د. يحيى عبد الرحيم

هيئة تحرير مجلة جامعة حلب في المناطق المحررة

رئيس هيئة التحرير: أ.د. أحمد بكار

نائب رئيس هيئة التحرير: أ.د. عماد برق

أعضاء هيئة تحرير البحوث التطبيقية	أعضاء هيئة تحرير البحوث الإنسانية والاجتماعية
أ.د. عبد العزيز الدغيم	أ.د. عبد القادر الشيخ
أ.د. ياسين خليفة	د. جهاد حجازي
أ.د. جواد أبو حطب	د. ضياء الدين القاش
أ.د. عبد الله حمادة	د. سهام عبد العزيز
أ.د. محمد نهاد كردية	د. ماجد عليوي
د. ياسر اليوسف	د. أحمد العمر
د. كمال بكور	د. محمد الحمادي
د. مازن السعود	د. عدنان مامو
د. عمر طوقاج	د. عامر المصطفى
د. محمد المجبل	د. أحمد أسامة نجار
د. مالك السلیمان	
د. عبد القادر غزال	
د. مرهف العبد الله	

أمين المجلة: هاني الحافظ

مجلة جامعة حلب في المناطق المحررة

مجلة علمية محكمة فصلية، تصدر باللغة العربية، تختص بنشر البحوث العلمية والدراسات الأكاديمية في مختلف التخصصات، تتوفر فيها شروط البحث العلمي في الإحاطة والاستقصاء ومنهج البحث العلمي وخطواته، وذلك على صعيدي العلوم الإنسانية والاجتماعية والعلوم الأساسية والتطبيقية.

رؤية المجلة:

تتطلع المجلة إلى الريادة والتميز في نشر الأبحاث العلمية.

رسالة المجلة:

الإسهام الفعّال في خدمة المجتمع من خلال نشر البحوث العلمية المحكمة وفق المعايير العلمية العالمية.

أهداف المجلة:

- نشر العلم والمعرفة في مختلف التخصصات العلمية.
- توطيد الشراكات العلمية والفكرية بين جامعة حلب في المناطق المحررة ومؤسسات المجتمع المحلي والدولي.
- أن تكون المجلة مرجعاً علمياً للباحثين في مختلف العلوم.

الرقم المعياري الدولي للمجلة ISSN: **2957-8108**

البريد الإلكتروني: journal@uoaleppo.net

الموقع الإلكتروني للمجلة: www.journal.uoaleppo.net

معايير النشر في المجلة:

- ١- تنشر المجلة الأبحاث والدراسات الأكاديمية في مختلف التخصصات العلمية باللغة العربية.
- ٢- تنشر المجلة البحوث التي تتوفر فيها الأصالة والابتكار، واتباع المنهجية السليمة، والتوثيق العلمي مع سلامة الفكر واللغة والأسلوب.
- ٣- تشترط المجلة أن يكون البحث أصيلاً وغير منشور أو مقدم لأي مجلة أخرى أو موقع آخر.
- ٤- يترجم عنوان البحث واسم الباحث والمشاركين أو المشرفين إن وجدوا إلى اللغتين التركية والإنكليزية.
- ٥- يرفق بالبحث ملخص عنه باللغات الثلاث العربية والإنكليزية والتركية على ألا يتجاوز ٢٠٠-٢٥٠ كلمة، وبخمس كلمات مفتاحية مترجمة.
- ٦- يلتزم الباحث بتوثيق المراجع والمصادر وفقاً لنظام جمعية علم النفس الأمريكية (APA7).
- ٧- يلتزم الباحث بألا يزيد البحث على ٢٠ صفحة.
- ٨- ترسل البحوث المقدمة لمحكمين متخصصين، ممن يشهد لهم بالنزاهة والكفاءة العلمية في تقييم الأبحاث، ويتم هذا بطريقة سرية، ويعرض البحث على محكم ثالث في حال رفضه أحد المحكمين.
- ٩- يلتزم الباحث بإجراء التعديلات المطلوبة خلال ١٥ يوماً.
- ١٠- يبلغ الباحث بقبول النشر أو الاعتذار عنه، ولا يعاد البحث إلى صاحبه إذا لم يقبل، ولا تقدم أسباب رفضه إلى الباحث.
- ١١- يحصل الباحث على وثيقة نشر تؤكد قبول بحثه للنشر بعد موافقة المحكمين عليه.
- ١٢- تعتبر الأبحاث المنشورة في المجلة عن آراء أصحابها، لا عن رأي المجلة، ولا تكون هيئة تحرير المجلة مسؤولة عنها.

جدول المحتوى

- ٩ أثر الذكاء الاستراتيجي في إدارة الأزمات
أ. فايز رشيد د. محمود عريض د. حسام خديجة
- ٤١ أخطاء زمن الفعل في الكتابات الامتحانية لدى طلاب اللغة الإنكليزية السوريين
أ. غياث الشيخ إبراهيم د. عبد الحميد معيكل
- ٦٩ الاستدلال بالمقاصد الشرعية في الثورة السورية
أ. عبد الرحمن اليوسف د. محمد راشد العمر
- ٩٥ الغزو السوفيتي لأفغانستان وأثره في التدخل الأمريكي المباشر في الخليج العربي
أ. أحمد ازعيتر د. سهام هنداوي
- ١١٩ القسم في مقام التبرئة
أ. جابر فرحان السلامة د. ضياء الدين عبد الغني القالاش
- ١٤٥ سلاح الغواصة ودور السلطان عبد الحميد الثاني في إدخالها إلى البحرية العثمانية
أ. حسام عبد المنعم الشحاد د. سهام محمد هنداوي د. عدنان مامو
- ١٦١ تَكَرُّرُ المَعْرِفَةِ فِي دِيْوَانِ (الأَعْشَى الكَبِيرِ)
أ. عدنان محمد الغزء د. ضياء الدين عبد الغني القالاش
- دراسة استقصائية لطرق التجزئة العميقة الخاضعة وغير الخاضعة للإشراف المستخدمة في استرجاع الصور
من قاعدة بيانات ضخمة
١٩٧
أ. صهيب بكور د. محمود موسى د. عبد الحافظ عبد الحافظ
- دراسة خصائص المياه الجوفية وصلاحيتها للاستخدامات المنزلية والزراعية لعينة مختارة من الآبار في
شمال محافظة حلب
٢٣٣
أ. علي شوبك د. بدر الدين الدخيل
- دور إدارة المواهب في الصحة التنظيمية لدى العاملين في جامعات الشمال السوري
٢٦٩
أ. حسام إبراهيم د. محمود عريض د. عمر درة
- مبدأ "المعاملة بالمثل" ومدى تطبيقه في الحالة السورية
٣١٥
أ. أيمن أحمد البيوش د. عبد الرحمن عزيزي
- علاقة الاغتراب النفسي بمعنى الحياة لدى عينة من طلاب كلية التربية في جامعة إدلب
٣٤٥
أ. باسم يوسف د. عبد الحي المحمود



دراسة استقصائية لطرق التجزئة العميقة الخاضعة وغير الخاضعة للإشراف
المستخدمة في استرجاع الصور من قاعدة بيانات ضخمة

إعداد

أ. صهيب بكور د. محمود موسى د. عبد الحافظ عبد الحافظ

ملخص البحث:

التجزئة هي إحدى الطرق الأكثر استخداماً لكفاءة الحوسبة والتخزين، فمع تطور التعلم العميق، تظهر طرق التجزئة العميقة مزايا أكثر من الطرق التقليدية من حيث البحث عن أقرب الجيران من أجل الحصول على العينات في قاعدة البيانات ذات المسافات الأصغر لعينات الاستعلام، وهي مهمة أساسية في كثيرٍ من المجالات، على سبيل المثال الرؤية الحاسوبية وتنقيب البيانات. وقد قمنا في هذه الدراسة الاستقصائية، بالدراسة التفصيلية لخوارزميات التجزئة العميقة الحالية بما في ذلك التجزئة العميقة من النوعين الخاضعة للإشراف وغير الخاضعة للإشراف، حيث نصنّف طرق التجزئة الخاضعة للإشراف العميق إلى طرق زوجية وطرق قائمة على الترتيب وطرق نقطية. وسوف نتحدث عن التكميم وفقاً لكيفية قياس أوجه التشابه بين أكواد التجزئة المكتسبة (learned hash codes). هذا ويتم تصنيف التجزئة العميقة غير الخاضعة للإشراف إلى أساليب قائمة على إعادة بناء التشابه وأساليب قائمة على الملصقات الزائفة وطرق قائمة على التعلم الذاتي خالية من التنبؤ بناء على أساليب التعلم الدلالية الخاصة بها، وفي الختام سنناقش بعض اتجاهات البحث المحتملة.

كلمات مفتاحية: التجزئة العميقة، استرجاع الصور، شبكات الطي العصبونية، التعلم العميق، تمثيل الصور.



A survey of the supervised and unsupervised deep hashing methods used to retrieve images from a large dataset

Prepared by:

Mr. Suhaib Bakour Dr. Mahmoud Mousa Dr. Abdel Hafez Abdel Hafez

Abstract

Hashing method is one of the most widely used methods of computing and storage efficiency. With the development of deep learning, deep hashing methods show advantages over traditional methods in terms of searching for nearest neighbors in order to get samples in the database with smaller query sample distances which is an essential task in many fields, for example computer vision and data mining. In this survey, we have detailed the existing deep hashing algorithms including both types of supervised and unsupervised deep hashing. We categorized deep supervised hashing methods into pairwise, ranking-based and pointwise methods. We will talk about quantization according to how similarities between learned hash codes are measured. Deep unsupervised deep hashing is categorized into similarity-reconstruction-based methods, pseudo-label-based methods and prediction-free self-learning-based methods based on their semantic learning manners. In conclusion we discuss some potential research directions.

Keywords: Deep Hashing, image retrieval, convolutional neural networks, deep learning, image representation.

١- مقدمة

يعدّ البحث عن أقرب الجيران (Nearest neighbor search) أحد المشكلات الأساسية في مختلف المجالات بما في ذلك الرؤية الحاسوبية والتعلم الآلي. والغرض منه هو العثور على أقرب نقطة من مجموعة البيانات إلى الاستعلام بناءً على مقياس مسافة معينة. ولكن عندما يكون لدينا بيانات ضخمة يكون التحدي في إيجاد العينة الأقرب إلى الاستعلام بدقة كبيرة وبتكلفة زمنية أقل ما يمكن. لمواجهة هذا التحدي، أولى الباحثون مؤخراً مزيداً من الاهتمام لتقريب البحث عن أقرب الجيران، لأنه في معظم الحالات يمكن أن يلبي احتياجات البحث، ويقلل على نحوٍ كبير من تعقيد البحث.

التجزئة هي إحدى الطرق الأكثر استخداماً لأنها فعالة جداً من حيث الحساب والتخزين [3]. والغرض منها هو تحويل نواقل الميزات عالية الأبعاد إلى أكواد تجزئة منخفضة الأبعاد، بحيث تكون أكواد التجزئة للكائنات المماثلة قريبة قدر الإمكان، وتكون أكواد التجزئة للكائنات غير المتشابهة مختلفة قدر الإمكان. تتكون طرق التجزئة الحالية من طرق التجزئة المحلية الحساسة وطرق تعلم التجزئة [5]. الغرض من التجزئة الحساسة المحلية هو تعيين البيانات الأصلية في عدة مجموعات تجزئة، وكلما كانت المسافة الأصلية بين العناصر أقرب، من خلال هذه الآلية، تم اقتراح عددٍ من الخوارزميات القائمة على التجزئة الحساسة محلياً [8]، والتي تظهر تفوقاً كبيراً في كل من الحساب والتخزين. ومع ذلك، من أجل تحسين معدل استدعاء البحث، تحتاج هذه الطرق عادةً إلى إنشاء عددٍ من جداول التجزئة المختلفة، لذلك لا تزال تطبيقاتها على مجموعات البيانات الكبيرة على نحوٍ خاصٍ محدودة.

مع تطور التعلم العميق، فإنّ الحصول على رموز التجزئة من خلال التعلم العميق يحظى باهتمام متزايد لسببين. السبب الأول هو أنّ قدرات التمثيل القوية للتعلم العميق يمكن أن تتعلم وظائف التجزئة المعقدة للغاية. السبب الثاني هو أنّ التعلم العميق يمكن أن يحقق أكواد تجزئة شاملة.

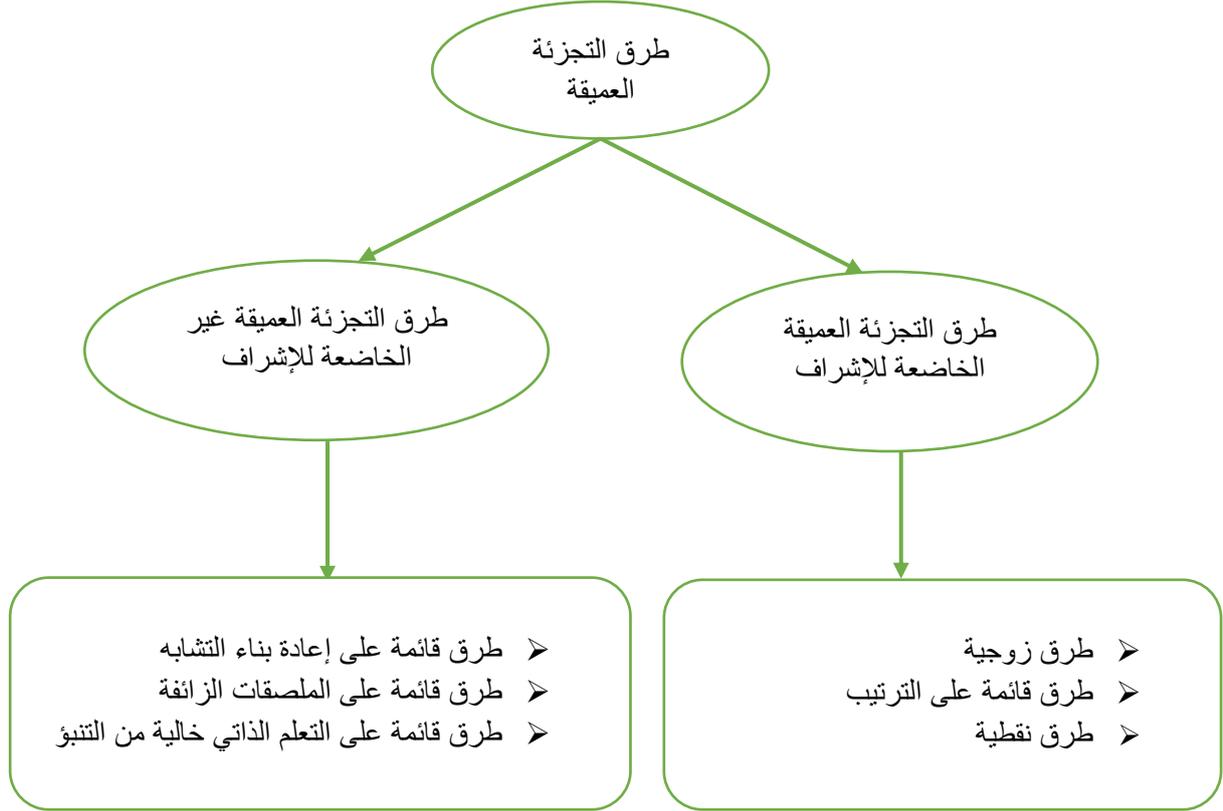
في هذه المقالة، نركز تركيزاً أساسياً على طرق التجزئة العميقة الخاضعة للإشراف وطرق التجزئة العميقة غير الخاضعة للإشراف، وهما عنصران رئيسان في بحث التجزئة. يشتمل تصميم طريقة التجزئة الخاضعة للإشراف العميق بنحوٍ أساسي على جزأين: تصميم هيكل الشبكة وتصميم دالة الخسارة. بالنسبة لمجموعات البيانات الصغيرة مثل [19] MINST و [16] CIFAR-10، يتم

استخدام المعمارية الضحلة مثل AlexNet[17] و CNN-F[18] على نطاق واسع. بينما بالنسبة لمجموعات البيانات المعقدة مثل NUSWIDE[7] وCOCO[22]، تحتاج بنية أعمق مثل VGG[29] وResNet50[35].

تم تصميم دالة الخسارة بقصد الحفاظ على هياكل التشابه. تهدف هذه الطرق عادةً إلى تضيق الاختلاف بين هياكل التشابه في الفضاء الأصلي وفضاء هامنغ [21]. وعادة ما يحصل الباحثون على أوجه التشابه في المساحة الأصلية باستخدام معلومات التسمية في سيناريوهات خاضعة للإشراف، تمت دراستها على نطاق واسع باستخدام طرق التجزئة العميقة المختلفة. ومن هنا فإنّ كيفية الحصول على أوجه التشابه بين رموز التجزئة المكتسبة مهمّ لخوارزميات مختلفة. نقوم أيضاً بتصنيف خوارزميات التجزئة الخاضعة للإشراف العميق وفقاً لكيفية قياس أوجه التشابه بين أكواد التجزئة المكتسبة إلى أربع فئات، أي الطرق الزوجية والطرق القائمة على الترتيب والطرق النقطية والقياس الكمي. لكل طريقة، كما نقوم بتحليل شامل لكيفية تصميم المقالات ذات الصلة بهدف التحسين والاستفادة من التسميات الدلالية، وكذلك الإضافات المستخدمة.

أما التجزئة العميقة غير الخاضعة للإشراف والتي لا تتطلب أية معلومات عن البيانات فإنها تلقى اهتماماً واسع النطاق مؤخراً نظراً لتطبيقها بسهولة في الممارسة العملية. في الإعدادات غير الخاضعة للإشراف، عادةً ما يتم اشتقاق المعلومات الدلالية من العلاقة الموجودة في المساحة الأصلية. استناداً إلى أساليب تعلم المعلومات الدلالية، نقوم بتصنيف خوارزميات التجزئة العميقة غير الخاضعة للإشراف إلى أساليب قائمة على الملصقات الزائفة، وطرق قائمة على إعادة بناء التشابه، وطرق قائمة على التعلم الذاتي خالية من التنبؤ، وسنتحدث عنها لاحقاً.

يظهر الشكل ١ الهيكل العام لهذه الدراسة حيث يظهر طرق التجزئة العميقة، والتي تنقسم إلى طرق التجزئة العميقة الخاضعة للإشراف، التي بدورها تنقسم إلى طرق زوجية وطرق قائمة على الترتيب وطرق نقطية، وطرق التجزئة العميقة غير الخاضعة للإشراف، التي تنقسم إلى طرق قائمة على إعادة بناء التشابه وطرق قائمة على الملصقات الزائفة وطرق قائمة على التعلم الذاتي خالية من التنبؤ. أخيراً نقدم أيضاً بعض مجموعات البيانات العامة شائعة الاستخدام ومخطط قياس أداء خوارزميات التجزئة العميقة. ثم إجراء مقارنة بين بعض الخوارزميات الرئيسية.



الشكل ١ الهيكل العام لهذه الدراسة الاستقصائية

مقارنةً بالمسوحات الأخرى حول طرق التجزئة [4]، تركز هذه المقالة البحثية على نحوٍ أساسي على طرق التجزئة العميقة الحديثة بدلاً من طرق التجزئة التقليدية وكيفية تحسينها لشبكة التجزئة، ودراسة كل من التجزئة العميقة الخاضعة للإشراف وكذلك التجزئة العميقة غير الخاضعة للإشراف على نطاق واسع.

٢- أساسيات:

٢-١ البحث عن أقرب جار nearest neighbor search:

يتمثل البحث عن أقرب جار في العثور على نقطة في مجموعة معينة الأقرب (أو الأكثر تشابهاً) إلى نقطة معينة، ويتم التعبير عنها عادة عن طريق دالة عدم التشابه، فكلما قل تشابه الكائنات زادت قيم دالة عدم التشابه.

بفرض لدينا مجموعة منتهية من النقاط $\Pi \subset \mathbb{R}^d$

$$NN(x_r) = \arg \min_{x_b \in \Pi} \rho(x_r, x_b) \quad (1)$$

حيث $x_r \in \mathbb{R}^d$ تشير إلى نقطة الاستعلام و ρ تشير إلى أي مقياس للمسافة مثل المسافة الإقليدية أو مسافة جيب التمام أو مسافة هامنج، وسنرمز للمسافة بنحوٍ عام بـ l_ρ .

طور الباحثون عدداً من طرق البحث الدقيقة عن أقرب الجيران [12]، والتي تعمل على نحو جيد عندما تكون d صغيرة. ومع ذلك، فإن البحث عن أقرب جار مكلف جداً بسبب تعددية الأبعاد [13] (curse of dimensionality)، وقد تم اقتراح سلسلة من الخوارزميات للبحث عن أقرب جار [15].

مبدأ هذه الطرق هو إيجاد أقرب نقطة ذات احتمالية عالية، بدلاً من إيجاد أقرب نقطة بدقة.

تتقسم خوارزميات البحث عن أقرب جار (Approximate Nearest Neighbors ANN) على نحو أساسي إلى ثلاث فئات: الطرق القائمة على التجزئة [25] والطرق القائمة على تكميم المنتج [34] والطرق القائمة على الرسم البياني [26]. حسنت هذه الخوارزميات على نحو كبير من كفاءة البحث مع ضمان دقة عالية نسبياً، لذا فهي مستخدمة على نطاق واسع في الصناعة. لكن الخوارزميات القائمة على التجزئة هي الأكثر دراسة من قبل الباحثين، لأنها تتمتع بإمكانيات كبيرة في تحسين كفاءة الحوسبة وتقليل تكلفة الذاكرة.

٢-٢ قياس التشابه ودوال التفعيل:

الرموز والمفاهيم الأساسية من أجل توضيح قياس التشابه ودوال التفعيل:

$$X = \{x_i\}_{i=1}^N \text{ تشير إلى المجموعة}$$

$$H = \{h_i\}_{i=1}^N \text{ تشير إلى خرج شبكة التجزئة}$$

$$h_i = \Psi\{x_i\}. B = \{b_i\}_{i=1}^N \text{ تشير إلى الرموز الثنائية التي تم الحصول عليها}$$

تشير إلى التشابه بين أزواج العناصر (x_i, x_j) في فضاء الدخل بـ S_{ij}^o وتشير إلى التشابه بين الزوجين (x_i, x_j) في فضاء هامنج بـ S_{ij}^{oh} .

فضاء الدخل يشير إلى مسافة الميزات المستخلصة من الصور على سبيل المثال المسافة الإقليدية

$$\|x_i - x_j\|_2, \text{ وأما قياس التشابه بين الأزواج يمكن حسابه عن طريق التابع الغاوسي مثلاً:}$$

$$\exp\left(-\frac{(d_{ij}^o)^2}{2\sigma^2}\right) \text{ و } I_{d_{ij}^o < \tau} \text{ حيث } \tau \text{ يمثل العتبة أو الحد ومن دوال قياس التشابه أيضاً التابع}$$

جيب التمام، وهو أكثر شيوعاً في التجزئة الخاضعة للإشراف العميق، حيث تكون القيمة 1 إذا كان

المثالان يشتركان في تسمية دلالية مشتركة و 0 خلاف ذلك. المسافة بين الزوجين d_{ij}^h في فضاء

هامنج تعرف كما يأتي:

$$d_{ij}^h = \sum_{l=1}^L \delta[b_i(l) \neq b_j(l)] \quad (2)$$

إذا تم تقييم قيمة ترميز التجزئة بـ 1 و 0 فلدينا:

$$d_{ij}^h = \|b_i - b_j\|_1 \quad (3)$$

أما إذا كانت قيم الترميز من 0 إلى L فيشار إلى التشابه في هذه الحالة:

$$S_{ij}^h = (L - d_{ij}^h)/L$$

أما إذا كان تقييم ترميز التجزئة بـ 1 و -1 فلدينا:

$$d_{ij}^h = \frac{1}{2}(L - b_i^T b_j) \quad (4)$$

ومن مقاييس التشابه أيضا الجداء الداخلي أي $S_{ij}^h = (b_i^T b_j + L)/2L$ ويمكن توسعتها إلى الحالة الموزنة كما يأتي:

$$d_{ij}^h = \sum_{l=1}^L \lambda_l \delta[b_i(l) \neq b_j(l)] \quad (5)$$

حيث يرتبط كل بت بوزن λ_l ، أما إذا كانت قيم الرموز 1 و -1 فلدينا:

$$d_{ij}^h = (b_i^T \Lambda b_j + tr(\Lambda))/2tr(\Lambda) \quad (6)$$

حيث $\Lambda = diag(\lambda_1, \lambda_2, \dots, \lambda_l)$ قطري و tr تشير إلى أثر المصفوفة، ويتم وضع كل بتات التجزئة المترابطة بشكل قطري في المصفوفة.

تعد دوال التفعيل ذات التصميم الجيد أحد أهم العوامل التي تؤثر على أداء التجزئة الخاضعة للإشراف العميق، حيث إن المبدأ الرئيس لتصميم دوال التفعيل هو الحفاظ على بنية التشابه، ما يعني تقليل الاختلاف بين أوجه التشابه في المساحات الأصلية ومساحات هامنج. نتيجة لذلك، تحتوي معظم دوال التفعيل على مصطلحات معلومات التشابه. من بينها، تكون دوال الخسارة النموذجية بطريقة زوجية، ما يجعل أزواج الصور المتشابهة لها رموز تجزئة متشابهة (مسافة هامنج صغيرة) والأزواج غير المتشابهة من الصور لها رموز تجزئة متباينة (مسافة هامنج كبيرة). وبعض الباحثين اعتمدوا على التشابه القائم على الترتيب للحفاظ على شروط الخسارة. على سبيل المثال، غالباً ما يتم استخدام الخسارة الثلاثية للحفاظ على أكبر قدر ممكن من الاتساق بين ترتيب عددٍ من العناصر المحسوبة من المساحات الأصلية ومساحات هامنج. هناك أيضاً عددٌ من مصطلحات الخسارة القائمة على القوائم التي تأخذ في الاعتبار مجموعات البيانات بأكملها للحفاظ على التشابه.

إلى جانب معلومات التشابه، يتم أيضاً استكشاف معلومات العلامات النقطية في تصميم دوال التفعيل على نحوٍ جيد. هناك ثلاث طرق شائعة للاستفادة من معلومات العلامات، وهي: الطريقة الأولى هي تراجع رموز التجزئة باستخدام العلامات. يتم ترميز العلامات في مصفوفة تنسيق واحد

ساخن one-hot وفقدان الانحدار، أي $\|Y - WH\|_F$ تضاف إلى دالة الخسارة. الطريقة الثانية هي إضافة طبقة تصنيف بعد شبكة التجزئة، ويتم إضافة خسارة التصنيف (على سبيل المثال، خسارة الإنتروبية التقاطعية) إلى دوال التفعيل. الطريقة الثالثة هو استخدام LabNet التي تم اقتراحها لأول مرة في 2018 [20]. يهدف LabNet إلى التقاط العلاقات الدلالية الواسعة بين أزواج الأمثلة. يستخدم مصطلح خسارة التكميم أيضاً استخداماً شائعاً في التجزئة الخاضعة للإشراف العميق، خاصة في طرق التجزئة القائمة على التكميم. الشكل النموذجي للتكميم هو معاقبة (penalize) المسافة بين الرموز المستمرة (أي مخرجات الشبكة) والرموز الثنائية. كأسلوب شائع في التجزئة العميقة، فإن فقدان توازن البت يعاقب الموقف الذي يكون لكل بت فرصة كبيرة لتكون 1 أو -1 بين مجموعة البيانات بأكملها ويمكن إضافة عددٍ من خسائر التنظيم إلى دالة الخسارة.

٣- خوارزميات أو طرق التجزئة العميقة لاسترجاع الصور:

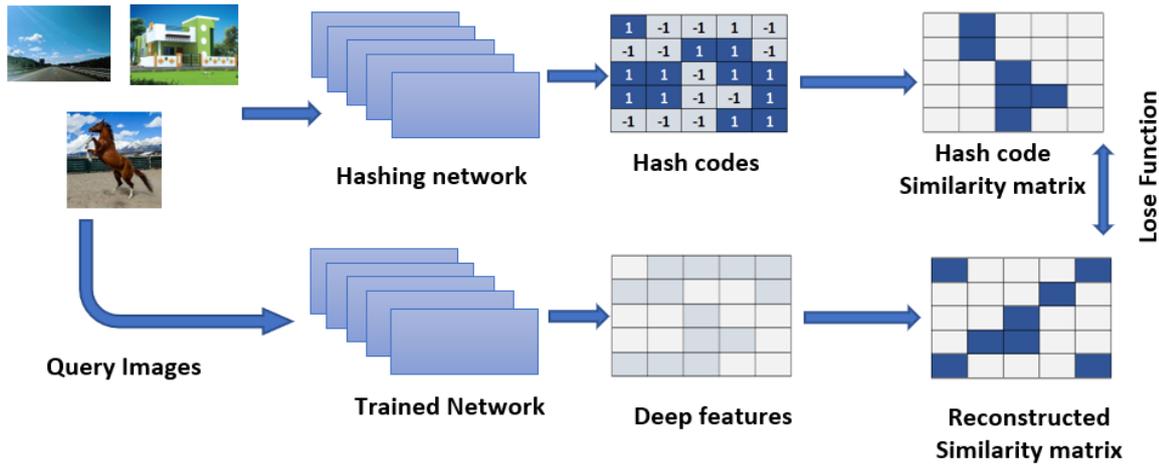
في هذه الدراسة الاستقصائية سنتحدث أولاً عن طرق التجزئة الخاضعة للإشراف العميق، والتي تشكل أساس تقنيات التجزئة العميقة غير الخاضعة للإشراف.

٣-١ نظرة عامة

تستخدم التجزئة الخاضعة للإشراف العميق الشبكات العصبية العميقة كدوال تجزئة، والتي يمكن أن تولد رموز التجزئة بطريقة شاملة، وبالتالي لدينا أربعة تحديات: ما بنية الشبكة العصبية العميقة المعتمدة؟ كيفية تصميم دالة الخسارة للحفاظ على بنية التشابه، كيفية تحسين الشبكة العصبية العميقة مع مشكلة التمييز، ما المهارات الأخرى التي يمكن استخدامها لتحسين الأداء؟ سنحاول في هذه الدراسة الاستقصائية الإجابة عن هذه التحديات.

٣-٢ طرق التجزئة العميقة الخاضعة للإشراف:

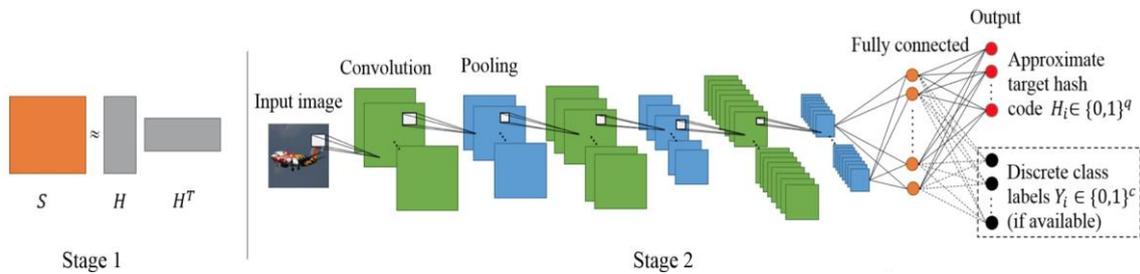
يوضح الشكل ٢ الإطار الأساسي للتجزئة الخاضعة للإشراف العميق مع قياس التشابه الزوجي. يتم إنتاج رموز التجزئة بوساطة شبكة تجزئة. بعد ذلك، تتم مطابقة معلومات التشابه الزوجي لرموز التجزئة والحقيقة الأساسية، وبالتالي الحصول على تشابه مع الحفاظ على الخسارة.



الشكل ٢ الإطار الأساسي للتجزئة الخاضعة للإشراف العميق، حيث يتم إنتاج رموز التجزئة بواسطة شبكة تجزئة ثم تتم مطابقة معلومات التشابه الزوجي لرموز التجزئة ثم الحصول على مصفوفة تشابه الحقيقة الأساسية، وبالتالي الحصول

٣-٢-١ التجزئة بشبكات الطي العصبونية CNNH

التجزئة بشبكات الطي العصبونية [2] CNNH هي الطريقة الأولى للاستفادة الكاملة من معمارية شبكات الطي العصبونية CNN. تعتمد هذه الطريقة على طريقة هبوط الإحداثيات المقيسة لتحليل مصفوفة التشابه الزوجية S إلى HH^T ، حيث H عبارة عن مصفوفة وكل سطر من أسطر هذه المصفوفة يمثل ترميز التجزئة التقريبي المرتبط بصورة التدريب. ثم م تعلم ميزات الصور التدريبية من خلال معمارية CNN بالإضافة إلى مجموعة من دوال التجزئة. وقد قام مكتشف هذه الطريقة بتحسين مرحلة تعلم دوال التجزئة من خلال دمج علامات فئة الصورة كجزء في طبقة الخرج. يوضح الشكل ٣ نظرة عامة على معمارية تجزئة شبكات الطي العصبونية CNNH، حيث يتكون من مرحلتين رئيسيتين، في المرحلة الأولى يتم حساب رموز التجزئة التقريبية عن طريق تقليل أخطاء إعادة البناء الآتية:



الشكل ٣ معمارية نموذج التجزئة الخاضعة للإشراف العميق [2] CHHN

$$\min_H \sum_{i=1}^n \sum_{j=1}^n (S_{i,j} - \frac{1}{q} H_i \cdot H_j^T)^2 = \min_H \|S - \frac{1}{q} HH^T\|_F^2 \quad (7)$$

حيث H عبارة عن مصفوفة بأبعاد $n \times q$ ، كل صف هو رمز تجزئة q - bit لصورة الهدف I_k

$$S_{i,j} = 1 \text{ عندما } I_i, I_j \text{ متشابه لغوياً}$$

$$S_{i,j} = -1 \text{ غير متشابه لغوياً}$$

في المرحلة الثانية، تتعلم تجزئة شبكات الطي العصبونية CNNH في وقت واحد مستخرج ميزة لصور الاستعلام إضافةً إلى مجموعة من دوال التجزئة، وفي هذه الحالة تستفيد شبكة الطي العصبونية من بنية الصور ثنائية الأبعاد ومعرفة الميزات الأساسية، ثم استخدام هذه الميزات للتحكم في نتائج رموز التجزئة المقابلة.

تجزئة شبكات الطي العصبونية CNNH مستنبطة من معمارية شبكات الطي العصبونية CNN، حيث يمكن عرض CNN على أنها مستخرج ميزة والنقاط البنية الدلالية الأساسية للصور، على الرغم من الاختلافات الكبيرة في المظهر، وبالتالي فهي تتفوق على معظم طرق التجزئة الكلاسيكية.

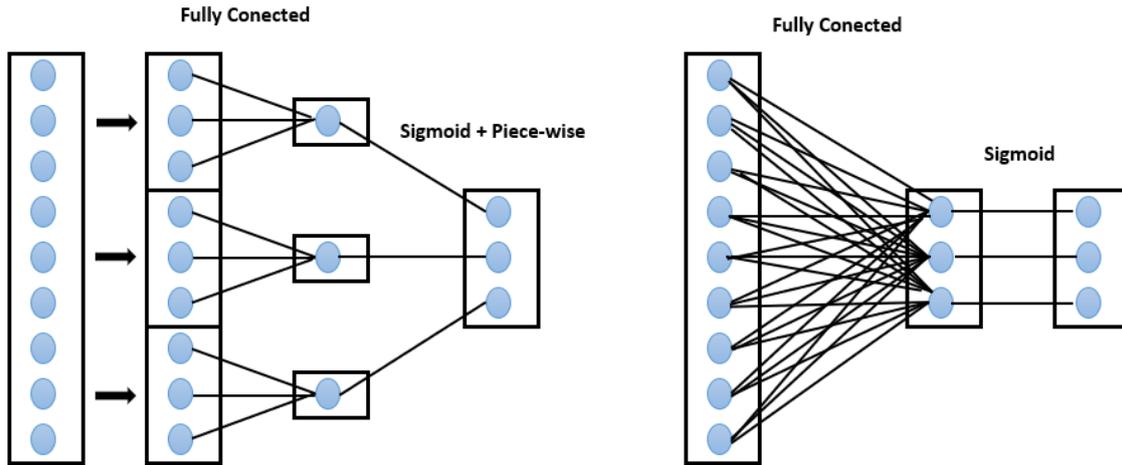
من عيوب هذه الطريقة المكونة من مرحلتين، يتم استخدام أكواد التجزئة التقريبية المكتسبة لتوجيه تعلم تمثيل الصورة، لكن تمثيل الصورة المكتسبة لا يمكنه تقديم ملاحظات لتعلم أكواد تجزئة تقريبية أفضل. يبدو أن هاتين المرحلتين في ترتيب عكسي بعض الشيء. نلاحظ أنه في التعلم الخاضع للإشراف، يمكن إنشاء مجموعة التشابه $S = \{S_{i,j}\}$ من العلامات الدلالية لنقاط البيانات أو التعليقات ذات الصلة في أنظمة الاسترجاع الحقيقية.

على الرغم من أن CNNH ليست طريقة شاملة، لا يمكن اعتبار متطلبات الحصول على مصفوفة التشابه عيباً رئيسياً، حيث يتم تضمين هذه المصفوفة كمعلومات خاضعة للإشراف في معظم الطرق الخاضعة للإشراف العميق.

٣-٢-٢ التجزئة الشبكة العصبية العميقة [1] (DNNH) Deep neural network hashing

في طريقة التجزئة لشبكات الطي العصبونية CNNH يتم تحديد رموز التجزئة التقريبية في بداية العملية بأكملها، والتي تبدو متناقضة مع منطقنا المشترك. وبالتالي قد لا يكون لزوج من الصور المتشابهة لغوياً أو غير المتشابهة متجهات مميزة بمسافة إقليدية صغيرة نسبياً أو كبيرة. هذا التفاعل أحادي الاتجاه محدود ويحتاج إلى التغيير. يولد تجزئة الشبكة العصبية العميقة (DNNH) [1] رموز تجزئة للصور من خلال معمارية شبكات الطي العصبونية CNN.

حيث يتم تحويل الصور المدخلة إلى تمثيلات صور موحدة عبر شبكة فرعية مشتركة من طبقات الطي المكدسة من خلال الشبكة العميقة، بعد ذلك، كما هو موضح في الشكل ٤، يتم ترميز تمثيلات الصور الوسيطة هذه في أكواد تجزئة عن طريق إجراء القسمة والترميز. حيث يتم تقسيم الميزات الوسيطة للإدخال إلى q شريحة، مع تعيين كل شريحة إلى بعد واحد بواسطة طبقة متصلة بالكامل. تم تصميم دالة خسارة الترتيب الثلاثي للحفاظ على أوجه التشابه النسبية وكذلك سحب الأزواج غير المتشابهة بعيداً.



الشكل ٤ طريقة تقسيم وترميز

في هذه الطريقة تم الاعتماد على معمارية الشبكة بحيث:

١- في نهاية الطبقة المتصلة بالكامل، يتم إنشاء كل بت تجزئة من شريحة منفصلة من الميزات لتقليل التكرار بين بتات التجزئة.

٢- من أجل تشجيع إخراج وحدة القسمة الترميز على أن تكون أكواد ثنائية فقد اعتمدت على دالة التنشيط السيني من خلال دالة عتبة متعددة التعريف.

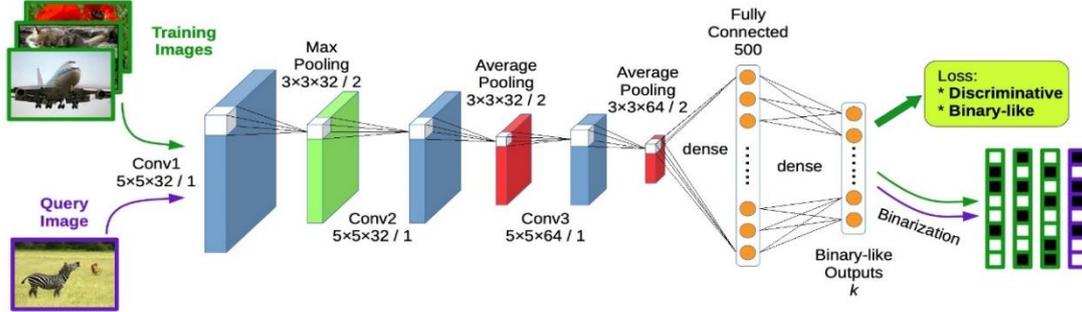
٣- تم تصميم دالة الخسارة الثلاثية لتمييز أن صورتين متشابهتين وأن الصورتين الأخرين غير متشابهتين. على عكس CNNH. هذه الطريقة أحادية الاتجاه ككل لمرحلتي التعلم وتؤكد من أ طبقة متصلة بالكامل وطبقة سينية البعض بحيث تتوافق عملية الترميز مع عملية استخراج الميزات.

٣-٢-٣ التجزئة الخاضعة للإشراف العميق [2] Deep supervised hashing (DSH)

تبنى طريقة التجزئة الخاضعة للإشراف العميق (DSH[2]) على معمارية شبكات الطي العصبونية CNN، حيث إن المدخلات عبارة عن زوج من الصور (بغض النظر عما إذا كانت

الصورتان متشابهتان)، والإخراج عبارة عن رمز ثنائي ويتم تصميم دالة الخسارة على نحو متقن، وذلك عن طريق التمييز في مساحة هامينغ عن طريق ترميز المعلومات الخاضعة للإشراف من صورة الإدخال، ثم يطبق تسوية على المخرجات ذات القيمة الحقيقية لتقريب القيم المنفصلة المرغوبة.

كما هو موضح في الشكل ٥



الشكل ٥ معمارية نموذج *DSH* [2] تتكون الشبكة من ٣ طبقات تجمع الالتفاف وطبقتين متصلتين بالكامل. تكون المرشحات في طبقات الالتفاف بحجم 5×5 مع الخطوة ١ (٣٢، ٣٢، ٦٤) مرشحاً في طبقات الالتفاف الثلاث على التوالي، طبقات تجميع بحجم 3×3 بخطوة ٢ (٣٢، ٣٢، ٦٤). تحتوي الطبقة الأولى المتصلة بالكامل على ٥٠٠ عقدة، والثانية (طبقة الإخراج) بها k عقدة (طول الكود). تم تصميم وظيفة الخسارة لتعلم الرموز الشبيهة بالثنائيات التي تحافظ على التشابه من خلال استغلال مصطلحات التمييز والمنظم. يتم الحصول على الرموز الثنائية عن طريق تحديد حجم مخرجات الشبكة للصور.

طريقة التجزئة الخاضعة للإشراف العميق (*DSH*) مبتكرة من نموذج *CNN* الذي يعالج أزواج الصور لتحديد ما إذا كانت الصورتان متشابهتين، ثم يولد رمز التجزئة المقابل بمسافة صغيرة نسبياً. تم تصميم دالة الخسارة للتأكد من توقع نقاط استعلام مماثلة لإغلاق النواتج في مساحة هامنغ مع إبعاد مخرجات النواتج غير المتشابهة بعيداً، حيث يتم تعريف دالة الخسارة على النحو الآتي:

$$L(b_1, b_2) = \frac{1}{2} (1 - y) D_h(b_1, b_2) + \frac{1}{2} y \max(m - D_h(b_1, b_2), 0) \quad (8)$$

$$s. t. \quad b_i \in \{+1, -1\}^k, j \in \{1, 2\}$$

بحيث $m > 0$ وسيط عتبة الهامش

الحد الأول من المعادلة يشير إلى الأزواج المتشابهة عند الإسقاط بعيداً (أي إبعادها بعيداً). الحد الثاني من المعادلة يشير إلى الأزواج غير المتشابهة المسقطه بالقرب عندما تنخفض مسافتهم في العتبة m

لنفترض أنه تم تدريب N زوج، يتم تعريف دالة الخسارة بالشكل الآتي:

$$L = \sum_{i=1}^N L(b_{i,1}, b_{i,2}, y_i) \quad (9)$$

$$s. t. b_{i,j} \in \{+1, -1\}^k, i \in \{1, \dots, N\}, j \in \{1, 2\}$$

هنا، لا يمكننا تحسين دالة الخسارة هذه مباشرةً بسبب القيود المنفصلة. الطريقة الشائعة هي استخدام دالة الظل (tanh) أو الدالة السينية (sigmoid) لتقريب القيود المنفصلة. ومع ذلك، فإن العمل مع هذه الدالة غير الخطية يؤدي إلى بطء في الشبكة.

يتم تصميم دالة الخسارة في طريقة التجزئة الخاضعة للإشراف العميق DSH منظم أو محسن

على النحو الآتي:

$$L(b_1, b_2) = \frac{1}{2}(1 - y) \|b_1 - b_2\|_2^2 + \frac{1}{2}y \max(m - \|b_1, b_2\|_2^2, 0) + \alpha(\|b_1\|_1 + \|b_2\|_1) \quad (10)$$

الفكرة الجديدة في هذه الطريقة هي فرض منظم أو محسن إضافي لاستبدال القيود الثنائية واستبدال الدالة السينية أو دالة الظل tanh، الدوال غير الخطية التي تقيد وتبقي تقارب الشبكة، خاصة عندما يكون الناتج بالقرب من الحدود. يتغلب المنظم أو المحسن المقترح على هذه المشكلة لأن تدرجه هو دائماً يأخذ القيمة ١

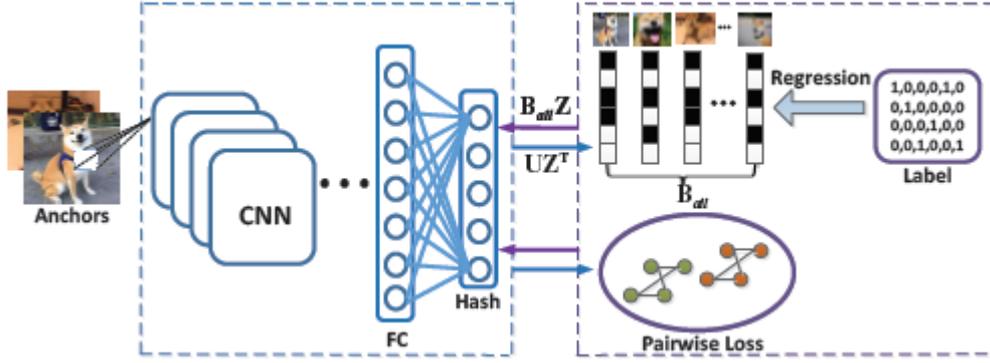
نلاحظ أن DSH عبارة عن بنية شاملة. تستخدم أكواد التجزئة التقريبية التي تم تعلمها لتوجيه

تعلم تمثيل الصورة ويمنح التمثيل الذي تم تعلمه ملاحظات لتعلم رموز التجزئة.

٣-٢-٤ تجزئة الرسم البياني ذات مركز العميق Deep Anchor Graph Hashing

(DAGH)[38]

جميع طرق التجزئة العميقة السابقة تعمل على إرجاع الميزات العميقة مباشرة إلى الرموز الثنائية المقابلة لتخفيف مشكلة التحسين المنفصلة. لذلك، في كل تكرار، لا يمكن تعلم سوى رموز ثنائية صغيرة دفعة واحدة، وبالتالي تحتاج هذه الأساليب إلى اعتماد عملية من خطوتين للحصول على الرموز الثنائية المثلى لمجموعة التدريب بأكملها، ما سيؤدي إلى فقدان المعلومات.



الشكل 6 معمارية نموذج DAGH [38] بين نظرة عامة. المربع الأزرق هو جزء من الشبكة العصبية والمربع الأرجواني هو جزء من دالة التفعيل. من خلال تصميم دالة الهدف القائمة على الرسم البياني المرسة، يمكن للإطار تحقيق الهدفين التاليين. من ناحية أخرى، يمكن استخدام خرج الشبكة ومصنوفة العلامات لإنشاء الرموز الثنائية لمجموعة التدريب ومن ناحية أخرى، يمكن للشبكة الحصول على تعليقات أكثر تمييزاً وفعالية من البنات المكتسبة والخسارة الزوجية.

في الرسم البياني للإرساء، يتم استخدام عدد ضئيل من نقاط الارتكاز لربط مجموعة البيانات بأكملها، ما يسمح بالحساب الضمني لأوجه التشابه بين الأمثلة المتميزة. في البداية، أخذ عينات من عدد من المراسي (المرتكزات) وبناء رسم بياني مرسة بين عينات التدريب والمراسي. ثم يمكن تقسيم دالة الخسارة إلى قسمين. يحتوي الجزء الأول على دالة خسارة احتمالية زوجية نموذجية ودالة خسارة انحدار خطي. في الجزء الثاني، يتم حساب الخسارة من خلال المسافات بين عينات التدريب والمراسي في نفس الفصل، ويتم استخدام كل من الميزات العميقة والرموز الثنائية لحساب المسافات. إلى جانب خسارة الاحتمالية الزوجية العامة وخسارة الانحدار الخطي، تقلل DAGH المسافات بين السمات العميقة لعينات التدريب والرموز الثنائية للمراسي التي تنتمي إلى نفس الفئة. تستخدم هذه الطريقة البيانات الموصوفة المتبقية على نحوٍ كامل في أثناء تدريب الدُفعات الصغيرة وتساعد في الحصول على أكواد تجزئة ثنائية فعالة.

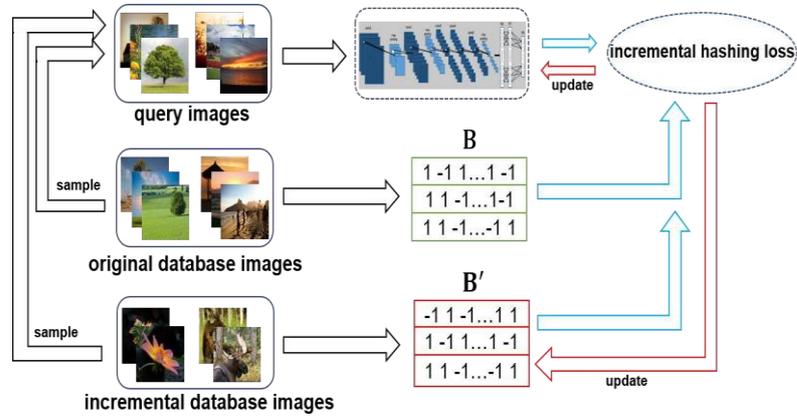
تشير التجارب التي أجريت على ثلاث مجموعات بيانات ضخمة إلى أن طريقة تجزئة الرسم البياني العميق DAGH تحقق أداءً استرجاع أفضل مع وقت تدريب أقل مقارنة بأساليب التجزئة العميقة السابقة.

3-2-5 شبكة التجزئة المتزايدة العميقة Deep Incremental Hashing Network (DIHN)[39]

تتطلب طرق التجزئة الخاضعة للإشراف قدرًا كبيراً من بيانات التدريب من فئات مختلفة. ومع ذلك، عندما تظهر صور فئات جديدة، يتعين على أساليب التجزئة العميقة السابقة إعادة تدريب

نموذج CNN وإنشاء رموز تجزئة لجميع صور قاعدة البيانات مرة أخرى، وهو أمر غير عملي لنظام الاسترجاع على نطاق واسع.

يوضح الشكل ٧ نظرة عامة على طريقة شبكة التجزئة المتزايدة العميقة DIHN



الشكل ٧ يظهر معمارية نموذج DIHN [39] حيث يتم استخدام نموذج CNN كدالة تجزئة فقط لاستعلام الصور، بينما يتم التعرف على رموز التجزئة الخاصة بصور قاعدة البيانات المتزايدة مباشرة. مع هذا التصميم غير المتماثل، يتم الاحتفاظ برموز التجزئة الخاصة بصور قاعدة البيانات الأصلية دون تغيير.

يمكن تلخيص المساهمة الرئيسية لهذه الطريقة على النحو الآتي:

- طريقة شبكة التجزئة المتزايدة العميقة DIHN هي أول نهج تجزئة عميق يمكنه أن يتعلم تدريجياً رموز التجزئة لتدريب الصور من فئات جديدة مع الاحتفاظ بالثبات الأصلي منها، حيث توفر الطريقة طريقة مرنة لتحديث نظام استرجاع الصور الحديثة.
- يتم وضع دالة فقدان التجزئة المتزايدة عن طريق الحفاظ على أوجه التشابه بين صور التدريب. وهو يدمج رموز التجزئة الثنائية الحالية للصور الأصلية لتدريب دالة التجزئة على صور الاستعلام. وفي الوقت نفسه، يمكن الحصول على أكواد التجزئة الثنائية للصور المتزايدة مباشرة أثناء التحسين أيضاً.
- تظهر التجارب المكثفة أن هذه الطريقة يمكن أن يقلل على نحو كبير من وقت التدريب مع عدم وجود خسارة تقريباً في دقة الاسترجاع، مقارنةً بأحدث الأساليب.

٣-٣ طرق التجزئة العميقة غير الخاضعة للإشراف DEEP UNSUPERVISED HASHING

٣-٣-١ مقدمة

عادة ما يكون لطرق التجزئة الخاضعة للإشراف للبيانات المصنفة أداء أفضل من خلال استخدام معلومات التسمية الدلالية. فإذا تمكنا أولاً من استخراج بعض معلومات التسمية الدلالية الخاضعة للإشراف من البيانات غير المصنفة ثم دمج هذه المعلومات الدلالية في عملية التدريب وبالتالي يمكن تحويل المشكلة إلى مشكلة خاضعة للإشراف. ومع ذلك فإن كيفية استنتاج معلومات التسمية الدلالية وكيفية استخدامها لتعلم رموز التجزئة هما مشكلتان رئيسيتان هنا.

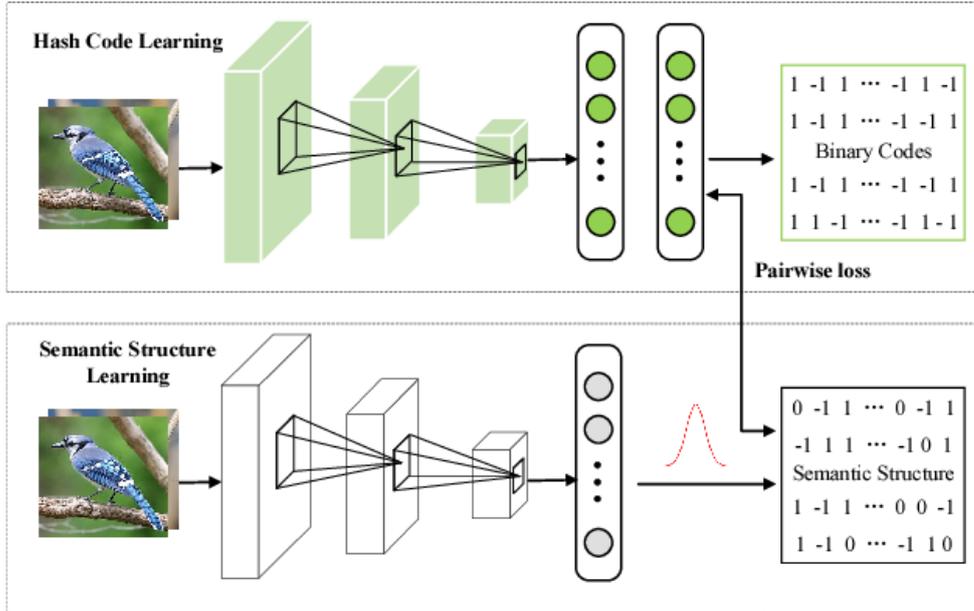
وفقاً لأدبيات تعلم الآلات، يمكن تصنيف الأساليب غير الخاضعة للإشراف على نحو أساسي إلى ثلاث فئات: الطرق القائمة على إعادة بناء التشابه، الطرق القائمة على العلامات الزائفة، الطرق القائمة على التعلم الذاتي الخالي من التنبؤ. وعادةً ما تولد الطرق القائمة على إعادة بناء التشابه معلومات دلالية زوجية ومن ثم الاستفادة من تقنيات الحفظ الدلالي الزوجية. وأما الطرق القائمة على العلامات الزائفة فتنتج تسميات زائفة نقطية للمدخلات ثم الاستفادة من تقنيات الحفظ الدلالي النقطية. أخيراً تستفيد الأساليب القائمة على التعلم والخاضعة للإشراف الذاتي الخالية من التنبؤ من البيانات نفسها للتدريب دون توليد معلومات دلالية صريحة، أي إشارات التشابه والتسميات الزائفة. نوضح بعض الطرق من هذه الفئات على النحو الآتي.

٣-٣-٢ التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف

(Semantic Structure-based Unsupervised Deep Hashing

SSDH)[33]

إن طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH من الطرق القائمة على إعادة بناء التشابه. حيث تم دراسة إحصائيات العلامات العميقة على نحو تجريبي، ووجد أن توزيع مسافة جيب التمام لأزواج النقاط يمكن تقديره بنصف توزيع غاوسي. لذلك تم بناء الهيكل الدلالي من خلال النظر في النقاط ذات المسافات الأصغر بوضوح من الآخرين على أنها متشابهة لغوياً والنقاط ذات المسافات الأكبر بوضوح من البقية باعتبارها متباينة لغوياً. ثم تم تصميم بنية عميقة ودالة خسارة زوجية للحفاظ على هذا الهيكل الدلالي في مساحة هامنج. الشكل ٨ يظهر إطار نموذجي عمل هذه الطريقة



الشكل ٨ معمارية نموذج SSDH[33]

طبقت طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH على نموذج VGG-F لاستخراج الميزات العميقة وأداء تعلم ترميز التجزئة ومسافة جيب التمام لكل زوج في مساحة الميزة العميقة، وقد وجد أن توزيع مسافات جيب التمام يمكن تقريبه بنصف توزيع غاوسي من خلال تقدير الوسائط.

طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH تحدد عتبتين للمسافة d_l, d_r وإنشاء بنية تشابه على النحو الآتي:

$$S_{ij}^o = \begin{cases} 1, & \text{if } d(z_i, z_j) \leq d_l \\ 0, & \text{if } d_l < d(z_i, z_j) < d_r \\ -1, & \text{if } d(z_i, z_j) \geq d_r \end{cases} \quad (11)$$

حيث d تشير إلى مسافة جيب التمام لمتجهين.

من المعادلة 11 تعدد طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH أزواج العينات التي تكون مسافة أصغر من متشابهة لغوياً بينما تعدد أزواج العينات ذات المسافات الأكبر متباينة لغوياً، وقد يتم اعتماد خسارة فرق التشابه على النحو الآتي:

$$L_{SSDH} = \sum_{i=1}^N \sum_{j=1}^N |S_{ij}^o| (S_{ij}^h - S_{ij}^o)^2 \quad (12)$$

حيث $S_{ij}^h = \frac{h_i^T h_j}{L}$, h_i تشير خرج الشبكة العميقة مع دالة تفعيل الظل بينما تم استخدام دالة التفعيل الإشارة في أثناء التقييم.

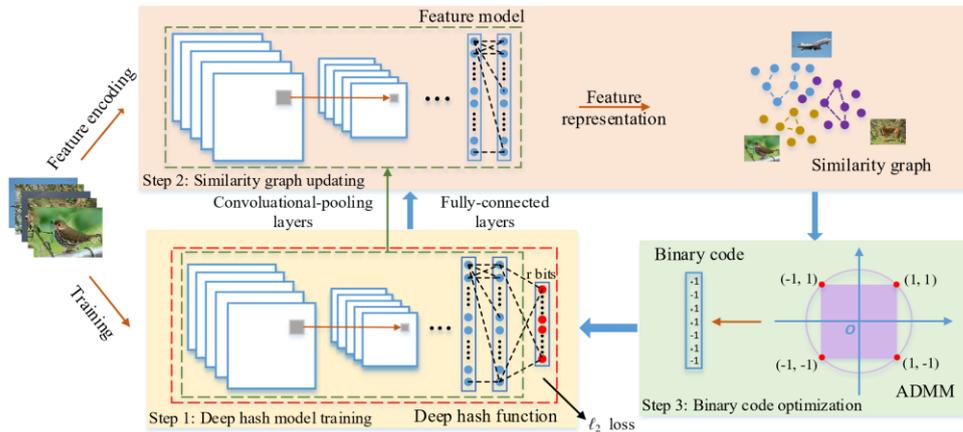
ومع ذلك، فإن أداء طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH محدود بسبب مشكلتين.

الأولى: لا يمكن الاعتماد على هيكل التشابه الخاص به باستخدام عتبتين تقريبيتين. الثانية: يتم تجاهل مجموعة من الإشارات في بنية التشابه. وقد أظهرت التجارب المكثفة أن طريقة التجزئة العميقة القائمة على البنية الدلالية غير الخاضعة للإشراف SSDH تتفوق تفوقاً كبيراً على أحدث الأساليب الحالية.

3-3-3 تجزئة التشابه - التكيفي العميق Similarity-Adaptive Deep Hashing (SADH)[40]

إن طريقة تجزئة التشابه - التكيفي العميق SADH [40] غير الخاضعة للإشراف من الطرق القائمة على إعادة بناء التشابه. في هذه الطريقة يتم بالتناوب أكثر من ثلاث وحدات تدريبية: تدريب نموذج التجزئة العميق، وتحديث الرسم البياني للتشابه، وتحسين الترميز الثنائي. يتمثل الاختلاف الرئيس عن طريقة التجزئة المكونة من خطوتين والمستخدم على نطاق واسع في أن تمثيلات الإخراج للنموذج العميق المكتسب تساعد في تحديث مصفوفة الرسم البياني للتشابه، والتي تُستخدم بعد ذلك لتحسين تحسين الكود اللاحق.

يبين الشكل 9 نظرة عامة على طريقة عمل تجزئة التشابه - التكيفي العميق SADH



الشكل 9 معمارية نموذج SADH [40] حيث تم تدريب نموذج التجزئة العميقة (في مستطيل الشرطة الحمراء) لتلائم الترميز الثنائي المكتسب. في الخطوة الأولى يتم أخذ جزء من النموذج المدرب لإنشاء تمثيل لوصفات ميزات الصور والتي تساعد في تحديث الرسم البياني للتشابه بين البيانات في الخطوة الثانية. باستخدام الرسم البياني للتشابه المحدث، يتم تحسين الترميز الثنائي بواسطة طريقة الاتجاه المتناوب للمضاعفات (ADMM) عبر النقاط

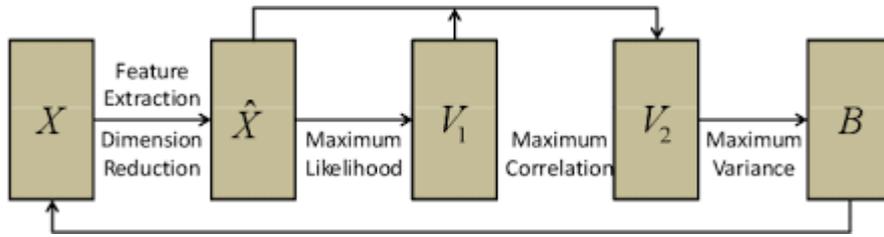
طريقة تجزئة التشابه - التكيفي العميق SADH تحافظ على التشابه من الصور الأولية كما هو موضح في الشكل ١ وتدعم طريقة تجزئة التشابه - التكيفي العميق SADH مجموعة كبيرة من دوال الخسارة، ولكن تم استخدام تجزئة الرسم البياني غير الخاضعة للإشراف. وقد أثبتت التجارب المكثفة فعالية SADH، التي تتفوق باستمرار على أحدث التقنيات من خلال وجود فجوات كبيرة.

٣-٣-٤ تجزئة عميقة غير خاضعة للإشراف مع علامات زائفة

Unsupervised Deep Hashing with Pseudo Labels

(UDHPL)[43]

في التجزئة العميقة غير الخاضعة للإشراف مع علامات زائفة UDHPL يتم أولاً باستخراج الميزات وتقليل أبعادها باستخدام خوارزمية تحليل المركبات الرئيسية (PCA) لتقليل الضوضاء. ثم يتم توليد العلامات الزائفة من خلال قاعدة بايز.



الشكل ١٠ معمارية نموذج [43] UDHPL حيث يتم استخدام رموز علامات زائفة المكتشفة للإشراف على ضبط شبكة التجزئة العميقة مثل الطرق الحالية الخاضعة للإشراف ويتم تنفيذ الإجراء الإجمالي T مره بنحو متكرر للعثور على نموذج شبكة

طريقة UDHPL قدمت مساهمتين رئيسيتين:

- تقوم بتحويل نموذج DH غير الخاضع للإشراف إلى الإشراف من خلال اكتشاف العلامات الزائفة.
- تعمل UDHPL على زيادة الارتباط بين متجهات الإسقاط للعلامات الزائفة والميزات العميقة، ويمكن عرض الميزات في مساحة هامنج.
- تعمل UDHPL على تحسين إنشاء كود التجزئة باستخدام مصفوفة التناوب.
- تعمل UDHPL على تحسين العلامات الزائفة من خلال PCA وتوجيه تدريب الشبكة من خلال تضخيم المعلومات المتبادلة، ما يساعد على الحفاظ على معلومات التشابه للاسترجاع الفعال.

٥-٣-٣ التعلم غير الخاضع للإشراف للصفات التمييزية والتمثيلات المرئية:

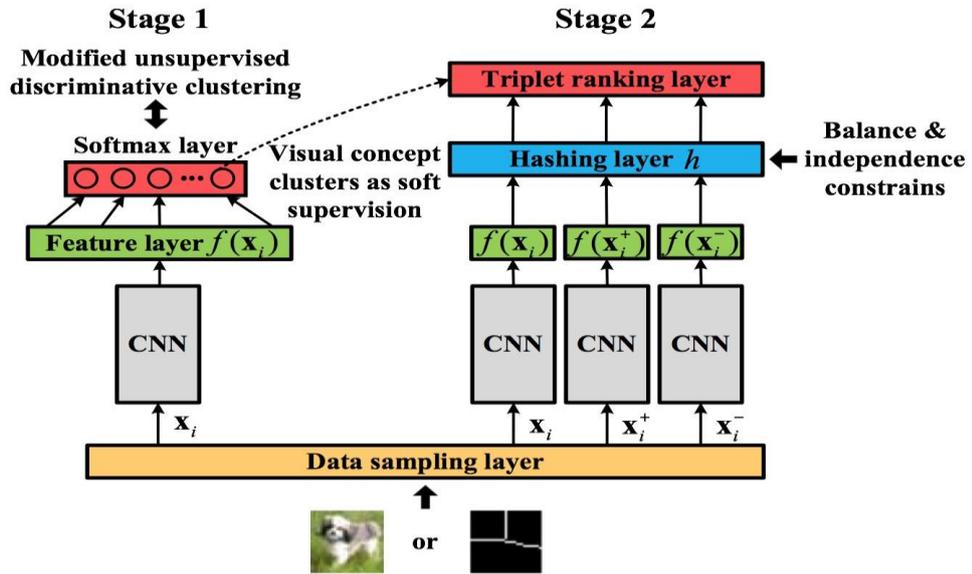
Unsupervised Learning of Discriminative Attributes and Visual Representations (DAVR)[14]

يتكون إطار عمل طريقة التعلم غير الخاضع للإشراف للصفات التمييزية والتمثيلات المرئية (DAVR[14]) كما هو موضح بالشكل ١١ من خطوتين.

(١) يتم تدريب شبكة CNN (جميع شبكات CNN لديها بنى وأوزان مشتركة) جنباً إلى جنب مع المجموعات التمييزية غير الخاضعة للإشراف [30] لإنشاء العنقدة (cluster membership) (مجموعة من البيانات المتشابهة).

(٢) يتم استخدام العنقدة (cluster membership) (مجموعة من البيانات المتشابهة) كأشرف ناعم لاكتشاف الصفات المشتركة من المجموعات (العناقيد) مع تحسين قابليتها لإمكانية فصلها باستخدام التفعيل الثلاثي.

تظهر الصفات التي تم تعلمها أنها قادرة على ترميز خصائص الصور الغنية من الصور الطبيعية. وبنحو عام، يتم تحويل التجزئة غير الخاضعة للإشراف إلى مشكلة خاضعة للإشراف بواسطة العلامات الزائفة التي تم الحصول عليها.



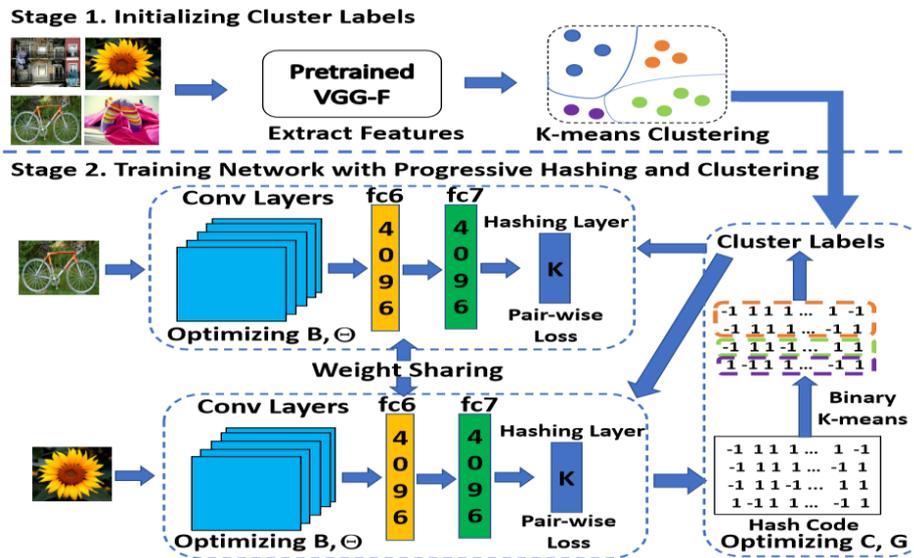
الشكل ١١ معمارية نموذج [14] DAVR تتكون طريقة DAVR من مرحلتين يتكون من مجموعات تمييزية غير خاضعة للإشراف وتجزئة خاضعة لإشراف الضعيف حيث يتم تعلم المجموعات المرئية ودوال التجزئة وتمثيلات الميزات بشكل مشترك.

٦-٣-٣ التجزئة العميقة بالمتوسطات K غير الخاضع للإشراف

Unsupervised Deep K-means Hashing (UDKH)[41]

تقوم التجزئة العميقة بالمتوسطات K غير الخاضع للإشراف (UDKH) بتجميع العقدة (التجميع) بالمتوسطات K. وتقوم هذه الطريقة باستخدام نتائج مجموعات K-means من أجل تهيئة علامات العقدة. التجزئة بالمتوسط العميق غير الخاضع للإشراف (UDKH) تتعلم كلاً من ترميز وعلامات العقدة بطريقة بديلة، حيث تقوم أولاً بإصلاح نتائج العقدة وتحسين ترميز التجزئة وكذلك شبكة التجزئة الخاضعة للإشراف ثم تقوم بإصلاح ترميز التجزئة والاستفادة من التصغير الخطي القريب المنفصل [28] Discrete Proximal Linearized Minimization لاشتقاق العلامات الزائفة المحدثة، ثم تقوم التجزئة بالمتوسط العميق غير الخاضع للإشراف (UDKH) بتكرار الخطوات السابقة حتى يتم التقارب. تعمل طريقة UDKH على تحسين جودة رموز التجزئة جنباً إلى جنب مع العلامات الزائفة من خلال التعلم التدريجي، ونظراً لأن رموز التجزئة المكتسبة تمييزية ما يحقق أداء أفضل مقارنة بالطرق غير الخاضعة للإشراف باستخدام العلامات الزائفة الثابتة. طريقة التجزئة بالمتوسط العميق غير الخاضع للإشراف (UDKH) تعمل أيضاً على تحسين دقة عقدة (تجميع) الصور وقد أظهرت التجارب على مجموعات بيانات الاختبار فعاليتها في استرجاع الصور وتجميعها.

يوضح الشكل ١٢ مراحل طريقة التجزئة العميقة بالمتوسطات K غير الخاضع للإشراف (UDKH).



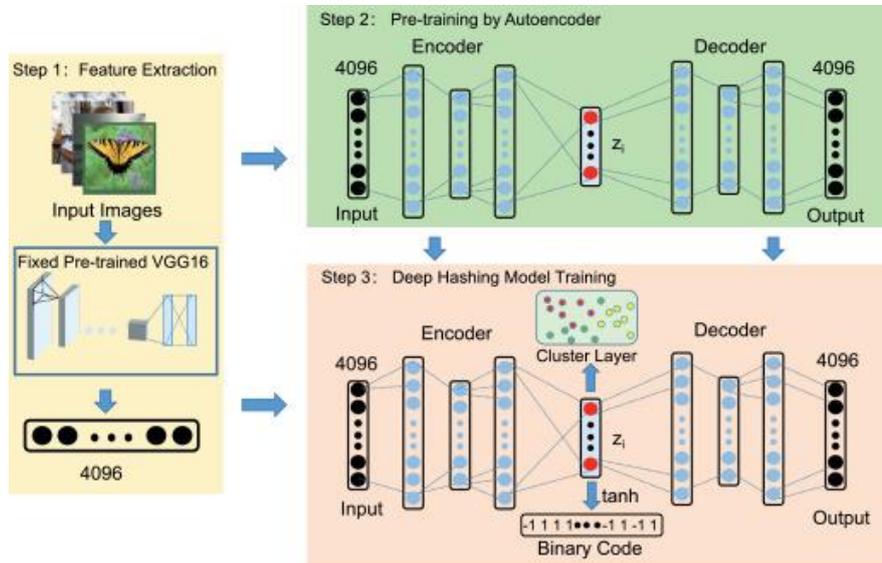
الشكل ١٢ معمارية نموذج UDKH[41]

٣-٣-٧ التجزئة العميقة المقادة بالعنقدة غير الخاضعة للإشراف:

Clustering-driven Unsupervised Deep Hashing (CUDH)[42]

تقوم طريقة التجزئة العميقة المقادة بالعنقدة غير الخاضعة للإشراف (CUDH) أولاً باستخراج الميزات العميقة من الشبكة المدربة مسبقاً، وهذه الطريقة مستنبطة من نموذج العنقدة العميق DEC [32] الذي يشكل العنقدة في مساحة التضمين ويعدل النموذج لتعلم العنقدة التمييزية على نحو متكرر في مساحة هامنج مع خسارة تكميم إضافية. التجزئة العميقة المقادة بالعنقدة غير الخاضعة للإشراف (CUDH) قادرة على إنشاء ترميز تجزئة تمييزي بحكم نموذج العنقدة العميق كما هو موضح بالشكل

١٣



يوضح الشكل ١٣ طريقة التجزئة العميقة المقادة بالعنقدة غير الخاضعة للإشراف [41] CUDH وهو عبارة عن ثلاث مراحل. ففي المرحلة الأولى استخراج الميزات من صور الدخل، والمرحلة الثانية مرحلة التدريب المسبق، والمرحلة الثالثة تدريب النموذج.

يمكن تلخيص مساهمة هذه الطريقة على النحو الآتي:

- CUDH عبارة عن شبكة عميقة شاملة (طرف لطرف end-to-end)، حيث يمكن تحسين جميع دوال التحسين بشكل مشترك باستخدام النسب المتدرج العشوائي (SGD) مع مجموعات صغيرة، ما يجعل بنحو عملي تحدي مجموعات البيانات واسعة النطاق.
- يمكن أن يحل مشكلة الحساب المسبق للعلامات الزائفة للطرق التقليدية، مثل طريقة k-means القائمة على أساس التجزئة وفرض قيود على مساحة العنقدة لإنتاج الترميز الثنائي مع الحفاظ على البنية المكانية.

وقد أظهر النتائج أن CUDH تتفوق على أحدث الأساليب بهوامش كبيرة.

٤- التقييم:

٤-١ مقاييس التقييم:

بالنسبة لخوارزميات التجزئة العميقة، تعتمد تكلفة الفضاء (space cost) على طول رموز التجزئة فقط، ولذلك عادة ما يتم الاحتفاظ بالطول كما هو عند المقارنة بين الخوارزميات المختلفة. يتم قياس كفاءة البحث من خلال متوسط وقت البحث للاستعلام، والذي يعتمد اعتماداً أساسياً على معمارية الشبكات العصبية. إلى جانب ذلك، إذا تم استخدام مسافة هامنج الموزنة، فلا يمكننا الاستفادة من العمليات على مستوى البت لتحقيق الكفاءة. كما نوقش أعلاه، عادة ما نستخدم دقة البحث لقياس الأداء. تشتمل المصفوفات الأكثر شيوعاً على متوسط الدقة والاستدعاء الدقة بالإضافة إلى منحنى الدقة - الاستدعاء. يتم تحديد الدقة من خلال نسبة العينات المرتجعة التي تشترك في العلامات العامة مع الاستعلام. يمكن صياغة الصيغة على النحو الآتي:

$$\text{precision} = \frac{TP}{TP+FP} \quad (13)$$

حيث TP تشير إلى عدد العينات التي تم إرجاعها، والتي لها علامات مشتركة مع الاستعلام FP تشير إلى عدد العينات التي تم إرجاعها، والتي ليس لها علامات مشتركة مع الاستعلام $precision @ k$ تعني أن العدد الإجمالي للعينات التي تم إرجاعها هو k أي $k = TP + FP$

الاستدعاء (Recall): يتم تعريف الاستدعاء بنسبة العينات في قاعدة البيانات التي لها علامات مشتركة مع الاستعلام الذي يتم استرداده.

يمكن صياغة الصيغة على النحو الآتي:

$$\text{recall} = \frac{TP}{TP+FN} \quad (14)$$

حيث FN هو العدد الإجمالي للعينات في قاعدة البيانات التي لها علامات مشتركة مع الاستعلام، بما في ذلك العينات التي لم يتم استرجاعها.

$recall @ k$ تعني أن العدد الإجمالي للعينات التي تم إرجاعها هو k .

منحنى استدعاء الدقة (Precision-recall curve): يتأثر كل من معدل الدقة ومعدل الاستدعاء في استرجاع الصورة بـ K . تتناسب معدلات الدقة والاستدعاء على نحوٍ عكسي. نتيجة لذلك، يمكننا إنشاء منحنى استدعاء الدقة عن طريق تغيير واستخدام معدل الدقة ومعدل الاستدعاء، على التوالي.

متوسط الدقة (Mean average precision (mAP)): عندما يكون الاختلاف في معدل الاستدعاء بين 0 و 1، يمكن حساب متوسط الدقة عن طريق تغيير معدل الدقة.

يتم استخدام نهج تجميع التسلسل لحساب متوسط الدقة في التطبيقات العملية مع التقدير:

$$AP = \frac{1}{F} \sum_{k=1}^N precision@K \Delta\{T@K\} \quad (15)$$

حيث $\Delta\{T@k\}$ تشير إلى التغيير في الاستدعاء من العنصر $k-1$ إلى k .

مجموع $\Delta\{T@k\}$ هو F والفكرة الأساسية لـ AP هي تقييم قائمة مرتبة عن طريق حساب متوسط الدقة في كل موضع. بعد ذلك، يمكن اشتقاق متوسط الدقة بأخذ متوسط الدقة لكل استعلام. في عديد من الأعمال، يتم حساب متوسط الدقة من حيث نتائج الاسترجاع ذات التصنيف K . يقوم بعض الباحثين أيضاً بحساب متوسط الدقة باستخدام Hamming Radius r ، عندما يتم النظر فقط في العينات التي لا تزيد مسافاتها عن r .

ألكسندر سابلايرولز وآخرون [135] أظهروا أن بروتوكولات التقييم الشائعة أعلاه للتجزئة الخاضعة للإشراف غير مرضية، لأن الحل المتواضع الذي يرمز ناتج المصنف يتفوق على نحو كبير على الطرق الحالية. علاوة على ذلك، فإنها توفر بروتوكول تقييم جيداً يعتمد على استرداد فئات غير المرئية ونقل التعلم. ومع ذلك، إذا كان تصميم طرق التجزئة يتجنب استخدام ترميز المصنف، فإن بروتوكولات التقييم الشائعة أعلاه لا تزال فعالة على نحو عام.

٤-٢ مجموعات البيانات (Datasets):

تتراوح مقاييس مجموعات بيانات التقييم المستخدمة من صغيرة إلى كبيرة إلى كبيرة جداً.

ويوجد نوعان لمجموعات البيانات مجموعات البيانات أحادية الوسم ومجموعات البيانات متعددة العلامات نعدد بعض من هذه الأنواع:

• MNIST [19] تتألف من 60,000 عينة تدريب و 10,000 عينة اختبار، وهي مجموعة بيانات ذات وسم واحدة، مصنفة إلى 10 فئات. يتم تمثيل كل صورة بـ 784-D ميزة خام (أولية) و 10,000 ميزة لاستعلام.

• CIFAR-10 [16] تتكون من 60,000 صورة من العالم الحقيقي في عشرة تصنيفات متميزة. وهي من مجموعة بيانات ذات وسم واحد، حيث تشير التصنيفات العشرة إلى الطائرات والسيارات

والطيور والقطط والغزلان والكلاب والضفادع والخيول والسفن والشاحنات. يتم تحديد هذه الأمثلة من خلال العلامات الدلالية المستخدمة لتقييم أداء طرق التجزئة المختلفة.

• ImageNet [9] عبارة عن مجموعة بيانات ضخمة تتكون من أكثر من 1.2 مليون صورة تم شرحها يدوياً بواسطة المشروع الضخم لمعرفة الأشياء المضمن داخل كل صورة، وهي مجموعة بيانات ذات وسم واحد وتتألف من 1000 فئة مثل "بالون" أو "قراولة".

• NUS-WIDE [7] هي مجموعة بيانات صور معروفة متعددة العلامات تم جمعها بواسطة فريق من NUS. وهو يتألف من 269648 مثالاً مع 5018 علامة فريدة. ترتبط هذه العينات يدوياً ببعض بـ 81 مفهوماً (مفاهيم)، ونظراً لأن الصور تحتوي على أكثر من وسم واحد، يتم التعامل مع عينتين على أنهما متشابهتان دلاليان إذا كانتا تشتركان في وسم دلالي واحد مشترك.

• MS COCO [22] هي مجموعة بيانات شائعة متعددة العلامات، تتكون من 82783 مثالاً تدريباً مع 4504 أمثلة للتحقق، كل منها مرتبط بجزء من الفئات الثمانية. بعد إزالة الأمثلة من دون أي معلومات عن الفئة، يمكن الحصول على 122218 عينة لتقييم أداء طرق التجزئة.

٣-٤ تحليل الأداء Performance Analysis :

٣-٤-١ مقارنة أداء الطرق الخاضعة للإشراف العميق:

نقدم نتائج بعض خوارزميات التجزئة والتكميم التمثيلية الخاضعة للإشراف العميق عبر CIFAR-10 و NUS-WIDE و ImageNet و MS COCO. بالنسبة لـ CIFAR-10، يتم اختيار 100 صورة عشوائياً لكل فئة (مما ينتج عنه 1000 صورة تماماً) كاستعلامات ويتم اعتماد باقي العينات كقاعدة بيانات. تشكل 500 عينة لكل فئة (مما ينتج عنه 5000 عينة بالكامل) مجموعة التدريب. بالنسبة لـ NUS-WIDE، تم اختيار مجموعة فرعية من 195834 عينة تتوافق مع 21 وسمًا، وهي الأكثر تكراراً (شيوعاً). بعد ذلك، يتم اختيار 100 عينة لكل فئة (مما ينتج عنه 2100 عينة بالكامل) كاستعلامات وتشكل العينات المتبقية مجموعة الاسترجاع.

يتم أخذ عينات من 500 صورة لكل فئة (مما ينتج عنه 10500 صورة تماماً) كمجموعة تدريب. بالنسبة إلى ImageNet، يتم اختيار 100 فئة عشوائياً. تشكل العينات المرتبطة بهذه الفئات في مجموعة التدريب قاعدة البيانات، ويتم استخدام العينات الموجودة في مجموعة التحقق كاستعلامات. يتم تحديد 100 صورة لكل فئة من قاعدة البيانات للتدريب. بالنسبة لـ MS COCO،

يتم استخدام ٥٠٠٠ عينة كاستعلامات، ويتم استخدام الباقي كقاعدة بيانات. تم اختيار ١٠٠٠٠ عينة من قاعدة البيانات للتدريب.

نلاحظ أنه بالنسبة للإعدادات التجريبية المختلفة، لم يتم عرض معظم النتائج التجريبية في هذا الملخص بالتفصيل. يتم عرض النتائج المقارنة لطرق التجزئة في الجدول ١ و ٢. من النتائج، هناك عددٌ من الملاحظات على النحو الآتي:

- تتفوق التجزئة الخاضع للإشراف بنحوٍ كبيرٍ على طرق التجزئة التقليدية (SDH) عموماً، ما يثبت القدرة القوية على التمثيل والتعلم للتعلم العميق.
- معلومات التشابه ضرورية للتجزئة العميقة. بالنسبة لأساليب التجزئة الخاضعة للإشراف العميق قبل عام ٢٠١٦، يتم الحصول على أكواد التجزئة في الغالب عن طريق نقل نماذج التصنيف دون معلومات التشابه الخاضعة للإشراف، بينما تتفوق الطرق ذات المعلومات الزوجية والترتيب عليها.
- تساعد معلومات التصنيف على زيادة أداء التجزئة العميقة. يمكن إظهار هذه النقطة من حقيقة أن DSDH يتفوق على DPSH بنحوٍ واضحٍ وتفوق LabNet. علاوةً على ذلك، تُظهر عددٌ من الأساليب النقطية (CSQ و OrthHash و PSLDH) أداءً مشابهاً مؤخراً من خلال تعيين العلامات في مساحة هامينج، ما يحقق أداءً رائعاً على مجموعات البيانات واسعة النطاق.
- تساعد عددٌ من المهارات بما في ذلك مصطلح التنظيم وتوازن البنات وتعلم المجموعات واستقلالية البنات في الحصول على أداء دقيق وقوي، وهو ما يمكن رؤيته من دراسات المستخلصة في بعض الأبحاث [29].
- على الرغم من أن طرق التجزئة الخاضعة للإشراف قد حققت أداءً رائعاً، إلا أنه من الصعب تطبيقها عملياً نظراً لأن التعليقات التوضيحية للبيانات واسعة النطاق لا يمكن تحملها. لمعالجة هذه المشكلة، توفر الأساليب غير الخاضعة للرقابة والقائمة على التعلم العميق حلاً فعالاً من حيث التكلفة لمزيد من التطبيقات العملية.

يظهر الجدول ١ مقارنة متوسط الدقة (mAP) على قاعدة البيانات CIFAR-10 و NUS-WIDE.

NUS-WIDE				CIFAR-10				الطرق Method
48bits	32bits	24bits	12bits	48bits	32bits	24bits	12bits	
0.608	0.625	0.618	0.611	0.522	0.509	0.511	0.439	CNNH[31]
0.715	0.713	0.697	0.674	0.581	0.558	0.566	0.552	DNNH[18]
0.748	0.731	0.731	0.712	0.799	0.770	0.742	0.644	DSH[23]

يظهر الجدول ٢ مقارنة متوسط الدقة (mAP) على قاعدة البيانات MS COCO و ImageNet.

MS COCO			ImageNet			الطرق Method
64bits	32bits	16bits	64bits	32bits	16bits	
0.567	0.574	0.564	0.553	0.449	0.281	CNNH[31]
0.609	0.603	0.593	0.565	0.460	0.290	DNNH[18]
-	-	-	0.665	0.550	0.348	DSH[23]

٤-٣-٢ مقارنة أداء الطرق العميقة غير الخاضعة للرقابة.

يعرض هذا الجزء نتائج طرق التجزئة العميقة غير الخاضعة للإشراف عبر CIFAR-10 و NUS-WIDE و MS COCO. نتبع الإعدادات في الأعمال السابقة [24، 27]. تقسيمات مجموعة البيانات للتدريب والاختبار وقاعدة البيانات هي نفسها في الفقرة 1-3-4 جزء من السجلات مقتبس من الأبحاث [27]. النتائج التي تمت مقارنتها موضحة في الجدول ٦. من النتائج، لدينا الملاحظات الآتية:

- تؤدي طرق التجزئة العميقة غير الخاضعة للإشراف أداءً أفضل بنحوٍ عام من الأساليب التقليدية (ITQ و AGH و DGH)، ما يشير إلى أن قدرة التعلم التمثيلية القوية للتعلم العميق مفيدة لأداء استرجاع الرموز الثنائية التي تم إنشاؤها.
- الأساليب التي تعتمد فقط مصطلحات التنظيم (DeepBits و UTH) تحصل على نتائج سيئة بين الطرق المقارنة، ما يدل على أن استكشاف المعلومات الدلالية أمر لا غنى عنه لأكواد

التجزئة التمييزية.

- الطرق التي تستكشف هياكل تشابه أكثر دقة (DATE و TBH) تتفوق في الأداء على الأساليب المبكرة التي تحصل على بنية التشابه بطريقة خشن (SSDH و DistillHash). السبب المحتمل هو أن إشارات التشابه الخاطئة ستؤدي إلى انتشار الخطأ في أثناء التعلم اللاحق لرمز التجزئة، ما يعني أداءً دون المستوى الأمثل.
- الأساليب التي تستخدم التعلم التباين (DATE و CIBHash) تحقق أداءً رائعاً بين الطرق المقارنة، ما يعني أن التعلم التباين هو أداة فعالة لتعلم رمز التجزئة التمييزي. مع تقدم البحث، يمكن لأساليب التعلم العميقة غير الخاضعة للإشراف أن تتفوق على جزء من الأساليب الخاضعة للإشراف العميق، وهو أمر مهم جداً.

يظهر الجدول ٣ مقارنة متوسط الدقة (mAP) للطرق غير الخاضعة للإشراف على قاعدة البيانات

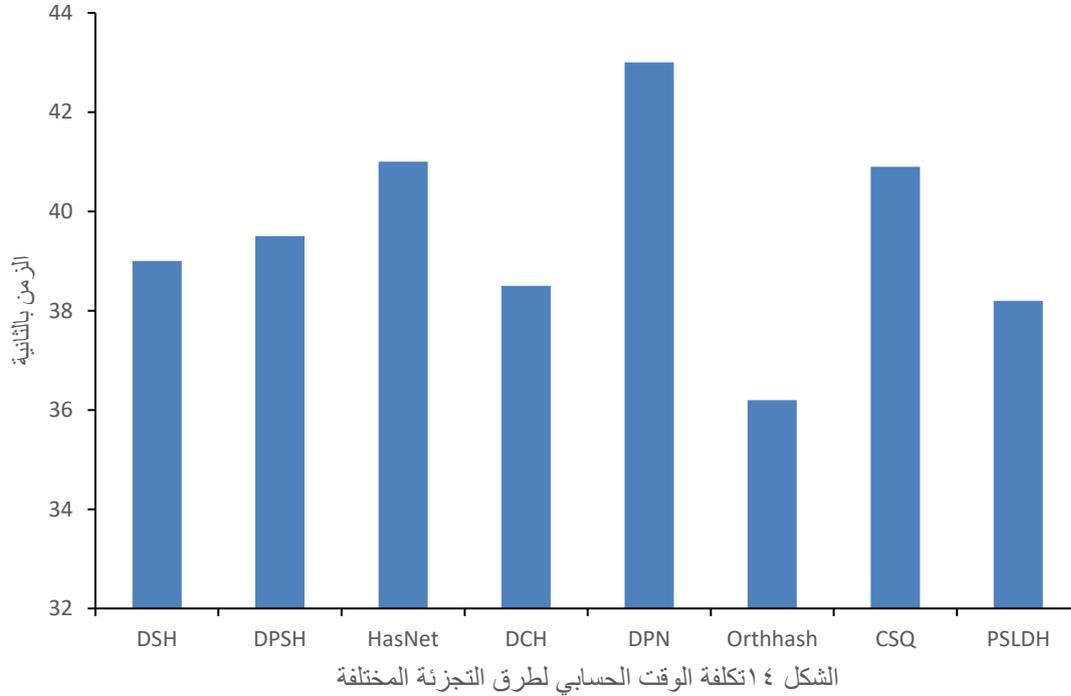
CIFAR-10, NUS-WIDE and MS COCO

MS COCO			NUS-WIDE			CIFAR-10			الطرق Method
64bits	32bits	16bits	64bits	32bits	16bits	64bits	32bits	16bits	
0.593	0.566	0.540	0.610	0.593	0.580	-	-	-	SSDH[33]
0.710	0.668	0.582	0.731	0.691	0.633	0.501	0.473	0.0.448	Greedy Hash[36]
0.508	0.465	0.438	0.549	0.495	0.450	0.324	0.307	0.287	UTH[37]

٥- تكلفة وقت التدريب

يظهر الشكل ١٤ كفاءة من التدريب لطرق التجزئة العميقة المختلفة. فقد اخترنا ثمانية طرق. يتم تحديد هذه الطرق بوساطة أعمدة أساسية مختلفة للشبكة (على سبيل المثال، AlexNet و VGG-F) ممثلة في المحور x . تم تحديد وسائط (بارامترات) جميع شبكات التجزئة بوساطة VGG-F ويتم تدريبها على وحدة معالجة رسومات NVIDIA GeForce GTX TITAN X. المحور y يظهر زمن التشغيل (الثانية) لكل فترة خلال مرحلة التدريب لمختلف الطرق. من النتائج، لدينا الملاحظات الآتية. أولاً: الاختلاف في الكفاءة بين هذه الأساليب محدود. السبب أن التكلفة الحسابية لطرق التجزئة

تعتمد بنحوٍ أساسي على إعادة التوجيه والانتشار الخلفي للعمود الفقري للشبكة. أساليب التحسين المحددة لها تأثيرات محدودة على التكلفة الحسابية. ثانياً: OrthHash هو الأكثر فاعلية بين الطرق المختلفة، وذلك لأن OrthHash يستفيد فقط من هدف واحد موجز في أثناء التحسين.



٦- التوصيات:

- ١- بناء نموذج تعلم عميق ذي فعالية حسابية عالية، ولا يتأثر بمتغيرات البيئة التي تطرأ على الصور من أجل تحسين أكثر لدقة الاستعلام عن الصور من قاعدة بيانات ضخمة.
- ٢- استخدام معايير تقييم إضافية لضمان جودة التقييم.
- ٣- العمل على تحديد الميزات المهمة المستخرجة من الصور من أجل تقليل العمليات الحسابية عبر تعليم موديل عصبوني عميق.
- ٤- الجمع بين خوارزميات البحث في أقرب الجوار التقريبي لتصميم طرق محددة أكثر للتجزئة العميقة.
- ٥- دمج تقنيات الشبكات العصبية العميقة المتطورة وتقنيات التعلم التمثيلي في التجزئة العميقة من أجل تعزيز تطوير استرجاع الصور على نطاق واسع.

٧- الخاتمة:

تم في هذه الدراسة الاستقصائية دراسة شاملة لأهم الأبحاث حول التجزئة العميقة الخاضعة للإشراف، والتجزئة العميقة غير الخاضعة للإشراف، وذلك استناداً إلى كيفية قياس أوجه التشابه بين رموز التجزئة، وقد تم تقسيم طرق التجزئة الخاضعة للإشراف العميق إلى ثلاث فئات: الطرق الزوجية، والطرق القائمة على الترتيب، والطرق النقطية. أما طرق التجزئة العميقة غير الخاضعة للإشراف فقد تم تقسيمها إلى ثلاث فئات بناء على طريق تعلم الدلالات، أي الطرق القائمة على إعادة بناء الشبكة والطرق القائمة على العلامات الزائفة والطرق القائمة على التعلم الذاتي الخالي من التنبؤ. الدافع الرئيس الذي نختاره لكيفية قياس أوجه التشابه بين رموز التجزئة المكتسبة للتصنيف هو أن الجوهر الأساسي لتعلم التجزئة هو الحفاظ على بنية التشابه حيث إن طرق قياس التشابه تحدد دوال الخسارة في التجزئة الخاضعة للإشراف العميق، إضافةً إلى ذلك العصبون.

نلاحظ أن طرق التجزئة العميقة الحالية تركز بنحوٍ أساسي على مجموعات البيانات العامة المصممة للتصنيف والكشف، والتي لا تعالج بنحوٍ كامل مشكلة البحث في أقرب الجيران.



المراجع

- [1] Beyer, K., Goldstein, J., Ramakrishnan, R., & Shaft, U. (1999, January). When is “nearest neighbor” meaningful?. In *International conference on database theory* (pp. 217-235). Springer, Berlin, Heidelberg.
- [2] Böhm, C., Berchtold, S., & Keim, D. A. (2001). Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. *ACM Computing Surveys (CSUR)*, 33(3), 322-373.
- [3] Cantini, R., Marozzo, F., Bruno, G., & Trunfio, P. (2021). Learning sentence-to-hashtags semantic mapping for hashtag recommendation on microblogs. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 16(2), 1-26.
- [4] Cao, Y., Qi, H., Zhou, W., Kato, J., Li, K., Liu, X., & Gui, J. (2017). Binary hashing for approximate nearest neighbor search on big data: A survey. *IEEE Access*, 6, 2039-2054.
- [5] Charikar, M. S. (2002, May). Similarity estimation techniques from rounding algorithms. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing* (pp. 380-388).
- [6] Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*.
- [7] Chua, T. S., Tang, J., Hong, R., Li, H., Luo, Z., & Zheng, Y. (2009, July). Nus-wide: a real-world web image database from national university of singapore. In *Proceedings of the ACM international conference on image and video retrieval* (pp. 1-9).
- [8] Dasgupta, A., Kumar, R., & Sarlós, T. (2011, August). Fast locality-sensitive hashing. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1073-1081).
- [9] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.
- [10] Do, T. T., Doan, A. D., & Cheung, N. M. (2016, October). Learning to hash with binary deep neural network. In *European Conference on Computer Vision* (pp. 219-234). Springer, Cham.
- [11] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T.,... & Houlsby, N. (2020). An image is worth 16x16



- words: *Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.*
- [12] Friedman, J. H., Bentley, J. L., & Finkel, R. A. (1977). *An algorithm for finding best matches in logarithmic expected time. ACM Transactions on Mathematical Software (TOMS), 3(3), 209-226.*
- [13] Fu, C., Song, L., Wu, X., Wang, G., & He, R. (2018). *Neurons merging layer: Towards progressive redundancy reduction for deep supervised hashing. arXiv preprint arXiv:1809.02302.*
- [14] Huang, C., Loy, C. C., & Tang, X. (2016). *Unsupervised learning of discriminative attributes and visual representations. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5175-5184).*
- [15] Jegou, H., Douze, M., & Schmid, C. (2010). *Product quantization for nearest neighbor search. IEEE transactions on pattern analysis and machine intelligence, 33(1), 117-128.*
- [16] Krizhevsky, A., & Hinton, G. (2009). *Learning multiple layers of features from tiny images.*
- [17] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks. NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 3-6 December. 1: 1097-1105.*
- [18] Lai, H., Pan, Y., Liu, Y., & Yan, S. (2015). *Simultaneous feature learning and hash coding with deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3270-3278).*
- [19] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). *Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.*
- [20] Li, N., Li, C., Deng, C., Liu, X., & Gao, X. (2018, July). *Deep joint semantic-embedding hashing. In IJCAI (pp. 2397-2403).*
- [21] Li, Y., & van Gemert, J. (2021, May). *Deep unsupervised image hashing by maximizing bit entropy. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 3, pp. 2002-2010).*
- [22] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D.,... & Zitnick, C. L. (2014, September). *Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.*



- [23] Liu, H., Wang, R., Shan, S., & Chen, X. (2016). *Deep supervised hashing for fast image retrieval*. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2064-2072).
- [24] Luo, X., Wu, D., Ma, Z., Chen, C., Deng, M., Huang, J., & Hua, X. S. (2021, October). *A Statistical Approach to Mining Semantic Similarity for Deep Unsupervised Hashing*. In *Proceedings of the 29th ACM International Conference on Multimedia* (pp. 4306-4314).
- [25] Lv, Q., Josephson, W., Wang, Z., Charikar, M., & Li, K. (2007, September). *Multi-probe LSH: efficient indexing for high-dimensional similarity search*. In *Proceedings of the 33rd international conference on Very large data bases* (pp. 950-961).
- [26] Malkov, Y. A., & Yashunin, D. A. (2018). *Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs*. *IEEE transactions on pattern analysis and machine intelligence*, 42(4), 824-836.
- [27] Qiu, Z., Su, Q., Ou, Z., Yu, J., & Chen, C. (2021). *Unsupervised hashing with contrastive information bottleneck*. *arXiv preprint arXiv:2105.06138*.
- [28] Shen, F., Zhou, X., Yang, Y., Song, J., Shen, H. T., & Tao, D. (2016). *A fast optimization method for general binary code learning*. *IEEE Transactions on Image Processing*, 25(12), 5610-5621.
- [29] Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. *arXiv preprint arXiv:1409.1556*.
- [30] Singh, S., Gupta, A., & Efros, A. A. (2012, October). *Unsupervised discovery of mid-level discriminative patches*. In *European Conference on Computer Vision* (pp. 73-86). Springer, Berlin, Heidelberg.
- [31] Xia, R., Pan, Y., Lai, H., Liu, C., & Yan, S. (2014, June). *Supervised hashing for image retrieval via image representation learning*. In *Twenty-eighth AAAI conference on artificial intelligence*.
- [32] Xie, J., Girshick, R., & Farhadi, A. (2016, June). *Unsupervised deep embedding for clustering analysis*. In *International conference on machine learning* (pp. 478-487). PMLR.
- [33] Yang, E., Deng, C., Liu, T., Liu, W., & Tao, D. (2018, July). *Semantic structure-based unsupervised deep hashing*. In *Proceedings of the 27th international joint conference on artificial intelligence* (pp. 1064-1070).

- [34] Zhang, T., Du, C., & Wang, J. (2014, June). *Composite quantization for approximate nearest neighbor search*. In *International Conference on Machine Learning* (pp. 838-846). PMLR.
- [35] He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [36] Su, S., Zhang, C., Han, K., & Tian, Y. (2018). *Greedy hash: Towards fast optimization for accurate hash coding in cnn*. *Advances in neural information processing systems*, 31.
- [37] Lin, K., Lu, J., Chen, C. S., & Zhou, J. (2016). *Learning compact binary descriptors with unsupervised deep neural networks*. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1183-1192).
- [38] Chen, Y., Lai, Z., Ding, Y., Lin, K., & Wong, W. K. (2019). *Deep supervised hashing with anchor graph*. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 9796-9804).
- [39] Wu, D., Dai, Q., Liu, J., Li, B., & Wang, W. (2019). *Deep incremental hashing network for efficient image retrieval*. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9069-9077).
- [40] Shen, F., Xu, Y., Liu, L., Yang, Y., Huang, Z., & Shen, H. T. (2018). *Unsupervised deep hashing with similarity-adaptive and discrete optimization*. *IEEE transactions on pattern analysis and machine intelligence*, 40(12), 3034-3044.
- [41] Dong, X., Liu, L., Zhu, L., Cheng, Z., & Zhang, H. (2020). *Unsupervised deep K-means hashing for efficient image retrieval and clustering*. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(8), 3266-3277.
- [42] Gu, Y., Wang, S., Zhang, H., Yao, Y., Yang, W., & Liu, L. (2019). *Clustering-driven unsupervised deep hashing for image retrieval*. *Neurocomputing*, 368, 114-123.
- [43] Zhang, H., Liu, L., Long, Y., & Shao, L. (2017). *Unsupervised deep hashing with pseudo labels for scalable image retrieval*. *IEEE Transactions on Image Processing*, 27(4), 1626-1638.