Triplet Network-Based DNA Encoding for Enhanced Similarity Image Retrieval

Takefumi Koike
Kyoto University, Japan
tkoike@easter.kuee.kyoto-u.ac.jp

Hiromitsu Awano
Kyoto University, Japan
awano@i.kyoto-u.ac.jp

Takashi Sato
Kyoto University, Japan
takashi@i.kyoto-u.ac.jp

ABSTRACT

With the exponential growth of digital data, DNA is emerging as an attractive medium for storage and computing. Thus, design methods for encoding, storing, and searching digital data within DNA storage are of utmost importance. This paper introduces image classification as a measurable task for evaluating the performance of DNA encoders in similar image searches. Furthermore, we propose a novel triplet network-based DNA encoder to improve the accuracy and efficiency. The evaluation using the CIFAR-100 dataset demonstrates that the proposed encoder outperforms existing encoders in retrieving similar images, with an accuracy of 0.77, which is equivalent to 94% of the practical upper limit, and 16 times faster training time.

ACM Reference Format:
Takefumi Koike, Hiromitsu Awano, and Takashi Sato. 2024. Triplet Network-Based DNA Encoding for Enhanced Similarity Image Retrieval. In *61st ACM/IEEE Design Automation Conference (DAC '24), June 23–27, 2024, San Francisco, CA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3649329.3657320

1 INTRODUCTION

Digital data is generated on a daily basis, and the annual global data generation is projected to surpass 160 Zeta Bytes in the near future [12]. Existing storage methods face limitations in terms of lifespan and power consumption. Deoxyribonucleic acid (DNA) storage emerges as a promising solution to these challenges.

DNA storage offers several advantages [16]. Firstly, it is capable of storing data at an extremely high density. It has been estimated that DNA can store 215 Peta Bytes per gram. Secondly, DNA serves as a reliable long-term storage medium. Conventional storage media, such as HDDs, magnetic tapes, and optical disks, are subject to data loss caused by accidental damage or natural corruption. In contrast, with proper management, DNA can be safely stored for thousands of years.Additionally, DNA storage is cost-effective due to its chemical stability and durability which eliminate the need for frequent replacements. Furthermore, it can be stored at room temperature, resulting in lower maintenance costs.

DNA can be utilized not only as a storage medium but also for a scalable computing platform. Building upon groundbreaking work in solving the traveling salesman problem [1], subsequent research has explored various applications, including solving satisfiability problems [9], data retrieval and random access [11] within DNA storage systems. A DNA-based similar image retrieval is proposed in [2]. This method transforms images into unique single-stranded DNA and evaluates image similarities through the degree of double-stranded DNA hybridization formed. An encoder facilitates this by leveraging a hybridization simulator to optimize DNA design. The retrieval is efficient due to its high level of parallelism, which enables scalable computational performance.

One challenge with existing similar image retrieval methods is the absence of quantitative performance comparisons. While earlier studies primarily focused on numerical similarity, it is also crucial to assess the visual similarity. Another challenge is the inefficiency of training DNA encoders. Constraints in simulation lead to prolonged training process, making it difficult to improve performance using scalable learning techniques.

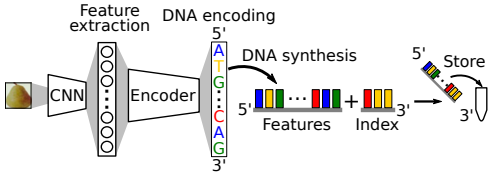
In this paper, we first propose a quantitative metric for evaluating the performance of DNA encoding methods in retrieving similar images based on visual similarity. We then propose a novel encoder network using deep metric learning [7, 13] for accurate and efficient similar image retrieval. The main contributions of this study can be summarized as follows:

- Introducing a novel method for quantifying accuracy in large-scale similar image retrieval using a classification framework, enabling quantitative comparison of accuracy among different methods.
- Proposing a highly accurate DNA encoder for similar image retrieval by integrating a triplet network. The integration of the deep metric learning successfully improves the accuracy of DNA-based image retrieval.
- Conducting comprehensive experiments using a large image dataset and CNN frameworks to compare the performances of similar image retrieval methods. The experiments with the CIFAR-100 dataset demonstrate that the proposed similar image retrieval method achieves an accuracy surpassing 94% of the software model while accelerating the training time by 16x.

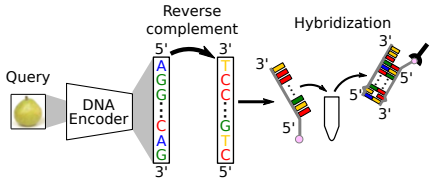
The remainder of this paper is structured as follows: Section 2 provides background information on DNA storage and image retrieval using DNA. Section 3 describes the proposed method for evaluating accuracy in image retrieval and proposes a novel DNA encoding method. Section 4 presents the numerical evaluation. Finally, Section 5 concludes this paper.

DAC '24, June 23–27, 2024, San Francisco, CA, USA

ACM ISBN 979-8-4007-0601-1/24/06
https://doi.org/10.1145/3649329.3657320



(a) Feature extraction and encoding



(b) DNA retrieval using DDH

Figure 1: CBIR using DNA. (a): Extract features from images and encode them into DNA using a DNA encoder. (b): Encode the query image into DNA and extract similar DNA sequences by hybridization.

2 BACKGROUND AND MOTIVATION

2.1 DNA and DNA hybridization

DNA is a nucleic acid that carries genetic information in living organisms. DNA is composed of deoxyribose, bases, and phosphate. A nucleoside is formed by a deoxyribose and a base, while a nucleotide is formed by a nucleoside and a phosphate. Phosphodiester bonds link the phosphate at the 5' position and the hydroxyl group at the 3' position, creating a single-stranded DNA. The phosphate end of the linked single-stranded DNA is referred to as the 5' end, while the end with the hydroxyl group of the deoxyribose is known as the 3' end. The sequence of nucleotide represents the order of nucleotide bases from the 5' end to the 3' end [10].

There are four types of bases in DNA: adenine (A), thymine (T), cytosine (C), and guanine (G). Each base pairs with a complementary base: A pairs with T, and C pairs with G through some hydrogen bonds. In double-stranded DNA, the bases form a double helix structure with the hydrogen-bonded bases facing inward and the phosphate backbone facing outward. This process of complementary bases pairing in single-stranded DNA sequence to form double-stranded DNA is known as DNA-DNA hybridization (DDH) [10].

Sequence similarity and stringency are important factors in selecting DNA pairs for DDH. Sequence similarity refers to the number of matching bases between the two nucleotide sequences. Higher similarity increases the likelihood of selecting them as a hybridization pair. Stringency, on the other hand, relates to the conditions of the hybridization process, such as temperature and pH. High stringency requires a strong sequence similarity between the DNA pairs, while low stringency conditions relax this requirement, allowing DDH even with some differences in the nucleotide sequences [4].

2.2 Content-Based Image Retrieval using DNA

Retrieving data from DNA storage is most efficiently accomplished utilizing DDH. Otherwise, all data must be decoded back into its original digital format. Accessing this data using CPUs and GPUs takes an impractically long time due to the immense amount of data involved. Methods that preserve the DNA structure are essential for successful data retrieval from a vast database.

The process of obtaining images based on local features is referred to as the Content-Based Image Retrieval (CBIR) task [18]. Literature [2] introduced a method for CBIR using DDH. Fig. 1(a) illustrates the conversion of feature representations from input images derived from a trained CNN model into DNA sequences. The feature vectors, extracted from a well-trained CNN, tend to be similar for related images [18]. By employing a DNA encoder trained to hybridize these vectors into similar DNA sequences, DNA mapping for CBIR is facilitated. Combined with an image index, the resulting encoded DNA forms a single-stranded DNA that is unique to each image.

The process for conducting searches using DDH is outlined in Fig. 1(b). Initially, the DNA assigned to the query image undergoes a transformation known as reverse complement. This operation reverses the sequence from the 5' to 3' end to the 3' to 5' end, with each base being substituted for its complementary counterpart. Subsequently, DNA is synthesized. The query DNA is amplified through Polymerase Chain Reaction (PCR), followed by hybridization in the DNA pool. The parallel nature of these processes efficiently identifies the DNA of similar images. Decoding the index representing the data portion of the extracted DNA enables the retrieval of similar images. This method is reported to achieve performance levels on a par with established benchmarks in approximate nearest neighbor algorithms.

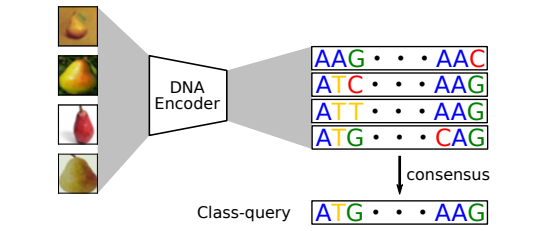
The existing methods have demonstrated similar image retrieval while maintaining DNA's structural integrity. However, the retrieval performance was evaluated by primarily comparing feature vector distances and hybridization yields, disregarding visual similarity. This evaluation, based purely on numerical values, consequently fails to provide a comprehensive assessment of DNA encoder performance. Furthermore, there are notable concerns regarding the encoder's training process. This process involves not only the encoder itself but also requires a CNN model, called the predictor, to predict hybridization. To enhance the encoder's performance, a sufficient amount of input data combinations that can represent the feature space must be prepared. However, the use of the predictor is inefficient as only one data can be compared at a time. In addition, hybridization prediction in the predictor utilizes convolutional and averaging layers for accurate inference, hence the training that involves both forward and back-propagation is time-consuming. Therefore, alternative approach are necessary.

In this work, we propose to assess DNA encoders' performance in terms of visual similarity through image classification using DDH. Additionally, we propose an accurate and efficient DNA encoder by introducing the idea of triplet network.

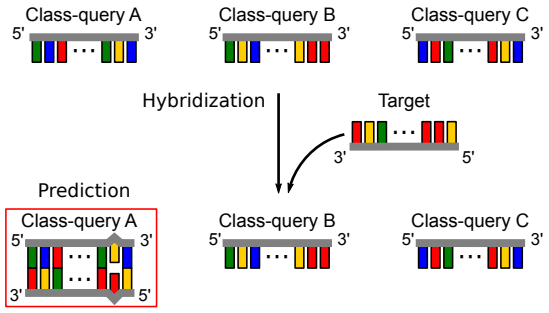
3 CLASSIFICATION VIA TRIPLET NETWORK

3.1 Image classification using DDH

To assess the performance of the CBIR methods objectively and quantitatively, we propose to utilize an image classification framework. The classification framework enables the comparison of methods using standard classification metrics, such as accuracy, recall,



(a) Class-query generation



(b) Class estimation

Figure 2: Image classification using DDH. (a) To determine a consensus sequence of each class query, the DNA sequence is generated by the most frequently occurring bases at each position. (b) Following the hybridization of the class query with the target data, the class of data is estimated based on the class query with the highest hybridization.

and precision. The lack of ground truth for image similarity has been a challenge in evaluating the performance of CBIR methods. Although mathematical metrics, such as Euclidean distance or cosine similarity, have been utilized, these metrics may not always match those perceived by human eyes and brain. In the classification framework, images belonging to the same class are considered to be more similar than those from different classes. Therefore, employing image classification is considered as an effective mean to quantitatively and objectively evaluate the performance of CBIR encoders.

To perform the classification task on the DNA sequences in the encoded dataset, a representative class-query must be defined for each class. These queries should be *mutually exclusive* and *collectively exhaustive* to ensure effective hybridization. As DNA sequences belonging to the same class share common base sequences and are close in terms of Hamming distance, we selected the representative query by identifying the most frequent base at each location within the class, extending from the 5' end to the 3' end, based on the training data (Fig. 2 (a)).

The encoder is trained to make the data distributions of the same class in the DNA space small and dense. To this end, we propose to apply the triplet network. Once the class-queries are determined, hybridization takes place between the query sequences and the target sequences. The query with the highest hybridization yield is predicted as the class to which the data belongs (Fig. 2 (b)).

3.2 Triplet network-based DNA encoding

To train a highly accurate DNA encoder, it is necessary to learn the similarity of the extracted features and reflect it in the similarity of the DNA sequences. Therefore, we utilize deep metric learning (DML), commonly employed for similarity learning of image features. Here, widely used pairwise training is not suitable to achieve high accuracy. Instead, simultaneously relating three subjects, i.e., bringing similar ones closer and dissimilar ones further in both the feature space and DNA sequences, is crucial for creating a reasonable feature space [7]. However, including hybridization yield estimation in the training pipeline can significantly lengthen overall training time due to the computationally intensive nature of DDH simulation. Therefore, we propose a new training method depending solely on the triplet network, which is based on the Hamming distance between DNA sequences instead of hybridization yield.

The outline of the proposed network is illustrated in Fig. 3. The proposed encoder network utilizes a triplet network to encode the same class of data into similar DNA sequences and other classes of data into the DNA sequences that are unlikely to hybridize.

The triplet network is a type of network used in DML. It accepts feature vectors as input and adjusts the distance of the data points in the output embedding space according to the class membership [7, 13]. As its name suggests, the triplet network operates based on the distances among three data points called triplets. The triplets consist of a set of data from the same class (referred to as "positive" or x_+), a set of data from a different class (referred to as "negative" or x_-), and a reference data point (referred to as "anchor" or x_0). The triplet network trains the machine learning model to minimize the distance between the anchor and positive data points while ensuring that the distance between the anchor and negative data points is greater than a predefined minimum distance called the margin.

During training, the weights are adjusted to map the anchor, positive, and negative data points into the embedding space. The triplet network allows for relative adjustment of the distance between classes. This relative adjustment is expected to enhance the separation of clusters in the embedding space.

In the embedding space, when the data points x_0, x_+, x_- are transformed to x'_0, x'_+, x'_-, respectively, the Triplet loss function is expressed using a distance metric, as follows:

$$\mathcal{L}(x'_0, x'_+, x'_-) = \max(0, d(x'_0, x'_+) - d(x'_0, x'_-) + m) \tag{1}$$

where m is the margin, and $d(\cdot)$ is an arbitrary distance function. In this work, the Hamming distance is used. As shown in Fig. 3, during the training, the encoder's outputs are extracted at their maximum values for each column to determine bases of each column. The Hamming distance, normalized to [0,1], between the anchor matrix $X'_0 = [x'_{00}\, x'_{01} \ldots x'_{0L}]$ and any data matrix $X' = [x'_0\, x'_1 \ldots x'_L]$ is expressed by:

$$d(X'_0, X') = \frac{1}{2L} \sum_{i=0}^{L} ||x'_i - x'_{0i}||_1 \tag{2}$$

where L is the length of a DNA and $||\cdot||_1$ is the L1-norm.

There are a large number of possible combinations for selecting three data points as the triplet. The choice of triplets is crucial to efficiently train an accurate model. In the proposed method, each

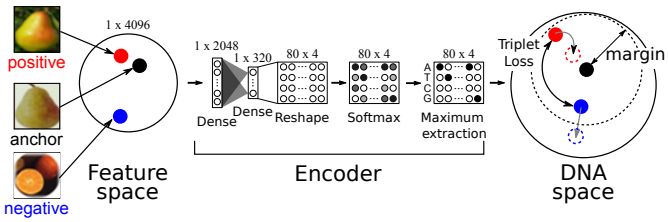


Figure 3: Proposed network structure. Features extracted by the CNN model are fed into the DNA encoder. Triplets are selected from the encoder output to utilize triplet loss in the training process.

batch of data, which serves as smaller unit of the training data, is assumed to asymptotically approximate the entire training data space, enabling time-efficient training.

Triplet selection [13] is performed as follows:

(1) Generate a distance matrix by calculating the distances between all combinations of data in the input batch, which is considerably smaller than the entire dataset.
(2) Choose an anchor data point and then classify positive and negative data points from the distance matrix based on their class labels.
(3) Randomly select data points from the data set that satisfy the following distance condition [13].

$$d(X_0', X_+') \leq d(X_0', X_-') < d(X_0', X_+') + m \tag{3}$$

By using triplets of data points that satisfy Eq. (3), the use of unnecessary data combinations can be avoided, resulting in efficient learning.

4 NUMERICAL EVALUATION

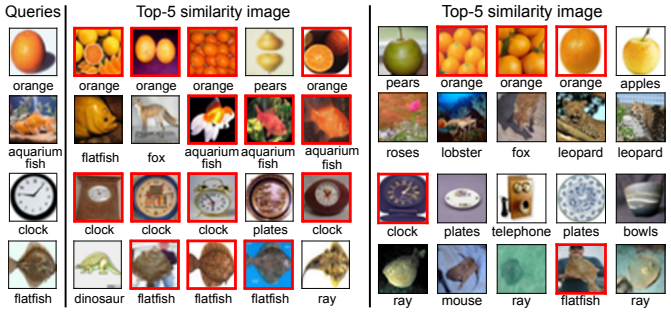
In this section, we compare the proposed method with the state-of-the-art method, primo [2], in terms of multi-class classification accuracy, the results of similar image retrieval using DDH, and the training time. For these tasks, we utilize CIFAR-10 and CIFAR-100 datasets [8]. Various CNN models [3, 6, 14, 15] were employed to investigate the impact of different feature spaces on the accuracy of the CBIR.

4.1 Evaluation setting

The image features we used in this section consist of a 4096 dimensional vector obtained from the intermidiate layer extracted by one of the CNN models. The training parameters are common for both the existing and proposed methods: a batch size of 128, 150 epoches, and 390 steps per epoch. In the proposed method, the distance between DNA sequences is represented as a normalized value in the [0, 1] interval, adjusted by the length of DNA sequence, and the margin of the triplet loss is fixed at $m = 0.8$. To assess the performance of the encoder, we employed NUPACK [5], a hybridization simulator, to mimic CBIR process. The simulator settings are summarized in Table 1.

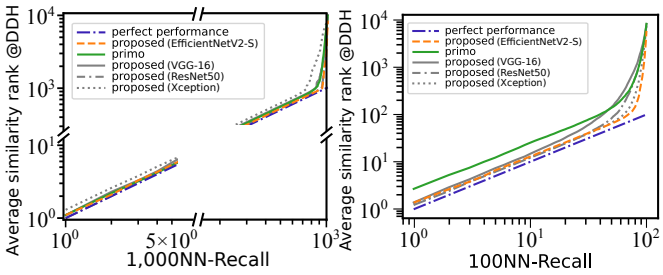
Table 1: NUPACK settings

Nucleotide length	Temperature	Target concentration	Query concentration
80 (nt)	21 (°C)	1.0 (nmol)	1.0 (nmol)



(a) Proposed method **(b) primo [2]**

Figure 4: Top-5 similar images retrieved using DDH with VGG-16 features (CIFAR-100). Images enclosed in red boxes belong to the same class as the query image.



(a) 1,000NN-Recall (CIFAR-10) **(b) 100NN-Recall (CIFAR-100)**

Figure 5: Average NN-Recall: The average number of images needed to select all classes of images (lower is better). Proposed methods approach optimal performance, highlighting proposed DNA encoder's nearly ideal coding capability.

4.2 CBIR results for different methods

Fig. 4 shows the top 5 similar images obtained using the DDH for four query images in class prediction with the VGG-16 features. In this example, there was no case where all the top 5 images belong to the same class label. However, despite the misclassification, the retrieved images still exhibit resemblances in color, shape, and object orientation. For instance, images misclassified as "orange" closely resemble "pears" in both shape and color. Similarly, images misclassified as "clock" retrieve "plates" with a similar shape. Additionally, images misclassified as "aquarium fish" and "flatfish" often exhibit a similar orientation to the query image, either facing left or right. Thus, even with the use of DDH for similar image search, the retrieved images capture the visual features of the query images.

We then evaluate whether image groups belonging to the same class is forming clusters in the DNA space. For this purpose, we

Table 2: Classification performance: The proposed method achieves equal or better performance than the state-of-the-art method across all indices. The top row shows the result of baseline CNN model, and the best accuracy is indicated in bold. The values in parenthesis represent the performance of the software CNN model.

CNN models	Methods	CIFAR-10 Accuracy	CIFAR-10 Recall	CIFAR-10 Precision	CIFAR-100 Accuracy	CIFAR-100 Recall	CIFAR-100 Precision
VGG-16	primo [2]	0.864 (0.854)	0.863 (0.854)	0.865 (0.858)	0.437 (0.631)	0.438 (0.631)	0.449 (0.657)
VGG-16	proposed	0.864 (0.854)	0.864 (0.854)	0.864 (0.858)	0.517 (0.631)	0.516 (0.631)	0.522 (0.657)
ResNet50	proposed	0.932 (0.927)	0.932 (0.927)	0.932 (0.928)	0.653 (0.712)	0.654 (0.711)	0.652 (0.717)
Xception	proposed	0.793 (0.783)	0.793 (0.783)	0.792 (0.785)	0.710 (0.749)	0.710 (0.749)	0.710 (0.756)
EfficientNetV2-S	proposed	**0.949** (0.947)	**0.949** (0.947)	**0.949** (0.947)	**0.766** (0.814)	**0.761** (0.814)	**0.765** (0.816)

Table 3: Training time comparison (CIFAR-100)

methods	Training time	
	Average (s/epoch)	Total (min.)
primo [2]	67.8	169
proposed	4.07	10.2

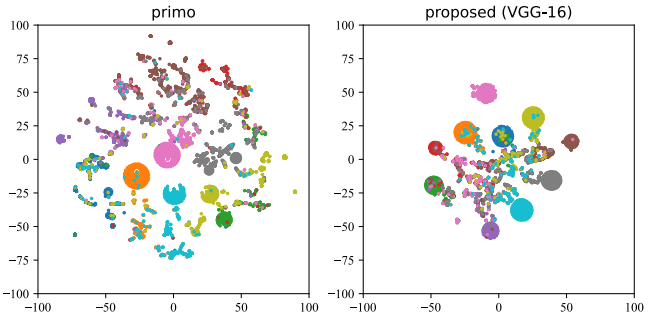


Figure 6: Visualization of DNA encoding space based on hybridization yields using t-SNE (CIFAR-10)

investigate the recall for each class and present the average as nearest neighbor recall (NN Recall) in Fig. 5. CIFAR-10's 1,000 NN Recall (Fig. 5a) demonstrates a performance very close to ideal. In datasets with a small number of classes and easily separable clusters due to high CNN accuracy, no significant difference is observed between primo and the proposed methods. However, as the number of classes increases, especially when dealing with complex datasets, the difference becomes prominent. CIFAR-100's 100 NN Recall (Fig. 5a) demonstrates that the proposed methods consistently outperform primo. The proposed methods again achieve nearly ideal performance, indicating excellent encoding ability in the DNA space for each class.

4.3 Classification accuracy

Owing to the adoption of the classification framework, the accuracy of the encoder can be estimated by evaluating its classification accuracy. As shown in Table 2, the classification accuracy using DDH varies significantly depending on the model. Interestingly, when classifying CIFAR-10, both the existing and proposed method achieve higher accuracy than the original software CNN model predictions. When classifying CIFAR-100, the accuracy of encoding and retrieval using DNA is lower compared to the original model. The proposed method consistently demonstrated equal or better classification accuracy than the existing method for all datasets.

The training time on a Quadro RTX 8000 with 4608 CUDA cores and 48 GB memory is summarized in Table 3. The proposed method significantly reduced the training time by more than 16x compared to primo.

4.4 Discussion

Figs. 4 and 5 show that the proposed method consistently retrieves a higher number of images from the same class, irrespective of the model's classification accuracy. To investigate this performance difference, the hybridization yields with each class query were visualized in Fig. 6 using t-SNE [17]. These plots indicate that the proposed method forms more distinct DNA clusters of the same class than primo, facilitating the retrieval of a significantly higher number of similar images. Additionally, our clusters spread widely in the DNA encoding space, in contrast to primo. In more intricate datasets like CIFAR-100, the clusters formed by the existing work tend to overlap, incurring erroneous image retrievals. These observations explain the remarkable performance that the proposed method achieves owing to the triplet-based training of the model. The proposed method properly maps image features to the DNA space, bringing similar features closer and dissimilar ones further apart, enabling better discrimination between different classes. It is worth noting that the proposed method achieved superior performance in shorter time, without necessitating computationally intensive hybridization predictors.

In Table 2, both primo and the proposed method demonstrated higher accuracy compared to the software CNN model on the CIFAR-10 dataset. Since both DDH methods utilized the same features extracted by the CNN model, the accuracy of the CNN model, operating purely in software, can be considered as the upper limit for each metric. Despite this, the accuracy improvement achieved through DDH-based methods is statistically significant, as confirmed by consistently higher accuracy in repeated experiments with different random seeds. The encoding process and nonlinear search enabled by DDH further reinforce the linear fully-connected network in the CNN model. Meanwhile, the experiments on the CIFAR-100 dataset did not reach the same level of accuracy as the CNN model. The difference in accuracy between CIFAR-10 and

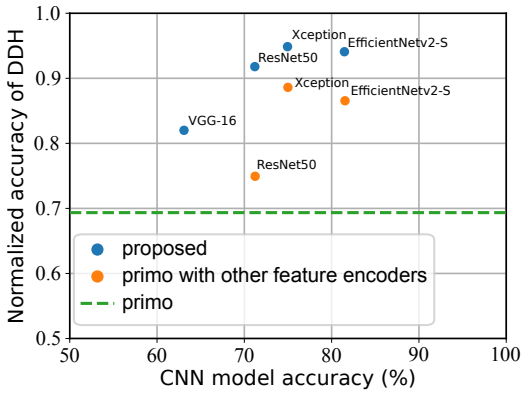


Figure 7: Classification performance normalized by the CNN model accuracy (CIFAR-100). DNA-based image retrieval accuracy increases with the CNN model accuracy. The proposed method outperforms the baseline with an performance of up to 94%.

CIFAR-100 can be attributed to the difference in CNN model accuracy. Dealing with a large number of classes calls for more sophisticated feature allocation in the encoding space. Addressing this discrepancy is one of our future works. However, we can confidently assert that the proposed method excels in capturing visual similarities. This assertion is substantiated by the comparison of image classification accuracy.

Fig. 7 shows the classification accuracy of CIFAR-100, normalized by that of the software CNN model, which serves as an upper limit, to highlight how the different training affects the performance. The comparison baseline is primo, and "primo with other feature encoders," in which feature encoders are replaced from the original implementation. Both Xception and EfficientNet achieved a performance higher than 94% (0.710/0.749 and 0.766/0.814, respectively, see Table 2). As the CNN model prediction achieves higher accuracy with more elaborated models, the accuracy of the classification using DDH approaches the upper limit defined by the CNN model. This result indicates that the proposed encoder performs exceptionally well when the quality of the input features improves. In addition, improving the CNN model used for feature extraction does not improve accuracy significantly when following the learning process of the existing work. The use of DML reflects the acquired knowledge of feature similarity. The proposed method demonstrates the potential to boost the accuracy of the DNA encoder by utilizing different CNN models and DML methods.

5 CONCLUSION

We proposed a novel DNA encoder utilizing a triplet network to enhance accuracy and computational efficiency in DNA-based similar image retrieval. Additionally, we proposed a quantitative evaluation method based on image classification to assess DNA encoder performance in similar image retrieval. Through experiments on large image datasets, the proposed DNA encoder, by incorporating DML methods into training and by simplifying hybridization yield prediction using Hamming distance, demonstrating a 16x faster training time and higher accuracy than the state-of-the-art DNA encoder. Moreover, based on the proposed class classification accuracy evaluation metric, our DNA encoder achieved a similarity image retrieval accuracy of 0.77, equivalent to 94% of the upper limit provided by the software-CNN on the CIFAR-100 dataset. The accurate and efficient retrieval of similar images through DNA technology demonstrates its potential for more complex operations.

ACKNOWLEDGEMENT

This work was partially supported by JSPS KAKENHI Grant Nos. 23K18462 and 23H03362.

REFERENCES

[1] Leonard M. Adleman. 1994. Molecular Computation of Solutions to Combinatorial Problems. *Science* 266, 5187 (1994), 1021–1024. arXiv:https://www.science.org/doi/pdf/10.1126/science.7973651
[2] Callista Bee, Yuan-Jyue Chen, Melissa Queen, David Ward, Xiaomeng Liu, Lee Organick, Georg Seelig, Karin Strauss, and Luis Ceze. 2021. Molecular-level similarity search brings computing to DNA data storage. *Nature Communications* 12 (2021), 4764.
[3] François Chollet. 2017. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1251–1258.
[4] J. H. Fisher, J. F. Gusella, and C. H. Scoggin. 1984. Molecular hybridization under conditions of high stringency permits cloned DNA segments containing reiterated DNA sequences to be assigned to specific chromosomal locations. *Proc. Natl Acad Sci USA* 81, 2 (Jan 1984), 520–524.
[5] Mark E. Fornace, Nicholas J. Porubsky, and Niles A. Pierce. 2020. A Unified Dynamic Programming Framework for the Analysis of Interacting Nucleic Acid Strands: Enhanced Models, Scalability, and Speed. *ACS Synthetic Biology* 9, 10 (2020), 2665–2678. https://doi.org/10.1021/acssynbio.9b00523 arXiv:https://doi.org/10.1021/acssynbio.9b00523 PMID: 32910644.
[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
[7] Elad Hoffer and Nir Ailon. 2015. Deep metric learning using triplet network. In *Similarity-Based Pattern Recognition: Third International Workshop*. 84–92.
[8] Alex Krizhevsky. 2009. Learning Multiple Layers of Features from Tiny Images. (2009), 32–33. https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf
[9] Richard J. Lipton. 1995. DNA Solution of Hard Computational Problems. *Science* 268, 5210 (1995), 542–545. https://doi.org/10.1126/science.7725098 arXiv:https://www.science.org/doi/pdf/10.1126/science.7725098
[10] D.A. Micklos. 2003. *DNA Science*. Cold spring harbor. https://books.google.co.jp/books?id=J9n-ngEACAAJ
[11] Lee Organick, Siena Dumas Ang, Yuan-Jyue Chen, Randolph Lopez, Sergey Yekhanin, Konstantin Makarychev, Miklos Z Racz, Govinda Kamath, Parikshit Gopalan, Bichlien Nguyen, et al. 2018. Random access in large-scale DNA data storage. *Nature biotechnology* 36, 3 (2018), 242–248.
[12] David Reinsel, John Gantz, and John Rydning. 2017. *Data Age 2025: The Evolution of Data to Life-Critical*. Technical Report. IDC.
[13] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 815–823.
[14] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
[15] Mingxing Tan and Quoc Le. 2021. Efficientnetv2: Smaller models and faster training. In *International conference on machine learning*. PMLR, 10096–10106.
[16] The DNA Data Storage Alliance. 2021. *Preserving our Digital Legacy: An Introduction to DNA Data Storage*. Technical Report.
[17] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 11 (2008).
[18] Ji Wan, Dayong Wang, Steven Chu Hong Hoi, Pengcheng Wu, Jianke Zhu, Yongdong Zhang, and Jintao Li. 2014. Deep Learning for Content-Based Image Retrieval: A Comprehensive Study. In *Proc. ACM Int. Conf. Multimedia*. 157.