**A New Algorithm for Sketch-Based Fashion Image Retrieval Based on Cross-Domain Transformation**

Summary of the Research

The article tackles the issues with traditional text and exemplar-based fashion image retrieval systems, introducing a novel algorithm employing cross-domain transformation for sketch-based fashion image retrieval.

Significance of Fashion Image Retrieval

Fashion image retrieval has gained importance with the surge of e-commerce platforms, and current methods using text or exemplar images face certain limitations.

Sketch-based image retrieval (SBIR) presents a more intuitive and natural method for users to convey their search needs. Nevertheless, the task of retrieving fashion images through sketches is complex due to significant cross-domain discrepancies.

This study introduces an inventive algorithm for sketch-based fashion image retrieval and provides a fine-grained sketch-based fashion image retrieval dataset.

Core Challenges in the Field

Different Domains Hurdle

There's a stark domain difference between fashion sketches and photos, which complicates the process of fine-grained fashion image retrieval.

Limitations of Current Retrieval Methods

Current methods relying on example images often clash with challenges such as subpar lighting, varying postures, and different angles of photography, which obstructs the retrieval accuracy.

Dataset Acquisition Issues

Amassing fashion sketches is daunting, as there's a scarcity of large-scale datasets dedicates to sketch-based fashion image retrieval.

The Three-Module Framework

The novel framework is constructed of three pivotal modules:

Cross-domain transformation module

Cross-domain feature extraction module

Cross-domain similarity measurement module

Key Research Contributions

Algorithm and Transformation

A new algorithm for sketch-based fashion image retrieval mightily bolstered by cross-domain transformation, which significantly enhances retrieval precision and masterfully addresses the heterogeneous nature of fashion sketches and photos.

Novel Retrieval Technique

A retrieval technique that ingenly transforms the fashion sketch query into a fashion photo, rummages through a fashion image dataset, and melds similarities across domains for heightened accuracy in results.

The Dataset Gift

The research also bears the fruit of a newly curated fine-grained sketch-based fashion image retrieval dataset, boasting 36,074 paired sketches and photos and spanning across 26 distinct fashion categories.

Insights into Related Works

SBIR Category and Detail Focus

Category-level SBIR zeroes in on fetching images from identical categories, whereas fine-grained SBIR zones in on the minutiae of the sketched object.

The use of CNN-based feature extraction has monumentally excelled the performances in sketch-based image retrievals.

The Scarcity in Fashion Datasets

An acknowledged impediment in sketch-based fashion image retrieval is the acute shortage of benchmark datasets.

The available datasets vary vastly in size and may be unimodal (solely photos) or multimodal (encompassing both sketches and photos).

GAN's Role in Advancements

Generative adversarial networks (GANs) have registered astounding breakthroughs in the realm of computer vision and foreshadow their utility in fashion image retrieval.

The Fruitful Outcome

The algorithm proposed, coupled with the detailed sketch-based fashion image retrieval dataset, confronts and surpasses the challenges faced by traditional retrieval methodologies, laying down an effective and trustworthy groundwork for future sketch-based fashion image retrieval.

Pointers to Cited References

For an in-depth understanding, consult the cited works and datasets listed in the article.

These notes offer a thorough encapsulation of the research article, preserving the quintessential aspects and all critical details as delineated in the document.

---# Fashion Image Retrieval using Cycle-Consistency GANs

GAN Mechanics

Generative Adversarial Network (GAN) consists of two modules: the generator (G) and the discriminator (D).

The generator aims to produce false images indistinguishable from real ones, while the discriminator aims to distinguish between real and false images.

GAN learning is a zero-sum game, where the discriminator aims to minimize its error in distinguishing between real and fake images, while the generator aims to maximize it.

The ideal condition is when it becomes difficult for the discriminator to distinguish between the two, i.e., D(G(z)) ≈ 0.5, where z is random noise.

Domain Gap Challenges

GANs are used to eliminate the domain gap between heterogeneous data such as sketches and photos.

The standard GAN requires paired training data, converting all sketches in the sketch domain to the same photo in the natural photo domain.

CycleGAN Improvements

CycleGAN, proposed by Zhu et al., allows bidirectional generation, transforming sketches into the photo domain, and vice versa, without the need for paired examples.

Methodology in Detail

The proposed method is based on the UNIT model and VGG-16 network.

It consists of three modules: cross-domain transformation, cross-domain feature extraction, and cross-domain similarity measurement.

The framework aims to retrieve fashion photos based on a query sketch.

Dataset Specifics

The dataset contains a wide range of fashion categories, including clothes, pants, skirts, and shoes.

It is a fine-grained dataset, with subcategories within each major category.

The dataset comprises 36,074 sketch-photo pairs, making it larger than existing datasets.

Sourcing Data

Photos were collected from online shopping websites, including TaoBao, Jingdong, and Amazon, along with a small portion from Baidu and Google.

The dataset includes specific types and styles of clothes, pants, skirts, and shoes.

Sketch Production

The collected photos were converted into corresponding sketches using the Structured Edge Detection Toolbox.

An erasing operation was performed on the edge maps to make them closer to free-hand sketches.

Transformation Techniques

A cross-domain transformation module was proposed, comprising fashion sketch and photo encoders, generators, and discriminators.

It includes two submodules: TS→P to transform sketches into photos and TP→S to transform photos into sketches.

The training process involves reconstructing the input and transforming data into the opposite domain, while ensuring similarity and minimizing the reconstruction error through variational autoencoders.

This detailed analysis provides a comprehensive understanding of the proposed method, dataset creation, and the implementation of GANs for fashion image retrieval. The method's focus on cross-domain transformation and feature extraction contributes to enhancing the retrieval accuracy, particularly in the context of fashion images.# Cross-domain Transformation and Feature Extraction for Fashion Image Retrieval

System Components Overview

The provided transcript discusses a comprehensive approach for cross-domain transformation and feature extraction for fashion image retrieval. The system is detailed and involves multiple submodules including Cross-Domain Transformation, Feature Extraction, and Similarity Measurement. The method involves using latent codes, VAE networks, discriminators, and cycle-consistency constraints to achieve accurate and robust transformation and retrieval of fashion images.

Transforming Fashion Photography

Encoder-Decoder Network for Fashion Photos

VAE Network: The fashion photo encoder (EncP) encodes input fashion photo (pi) into a latent code (zP), and the fashion photo generator (GenP) decodes the latent code to reconstruct the fashion photo.

Objective Function (LEncP): Defines the encode-decode process by penalizing latent code distribution deviation and constraining the reconstructed photo to be similar to the input photo.

Authenticity Discrimination for Fashion Photos

The discriminator DisP distinguishes between real fashion photos and generated fashion photos (from GenP).

Objective Function (LGANP): Comprises terms to differentiate between real and generated fashion photos within the GANP composed of GenP and DisP.

Sketch Generation and Validation

Transition to Fashion Sketches

Fashion sketch generator (GenS) and discriminator (DisS) constitute GANP→S, used for transforming fashion photos to fashion sketches and distinguishing between real and transformed sketches.

Preserving Originality via Cycle-Consistency

Ensures that transformed fashion photos can be accurately reconstructed back to the original photo, and original sketch features are not lost after transformation.

Uses cycle-consistency constraints for both fashion sketch and photo transformations.

Combining Objectives for Transformation

Total objective functions for the fashion sketch and photo cross-domain transformation submodules are given, comprising the respective individual objective functions.

Optimal Training and Testing Methods

Crafting the Training Procedure

Utilizes Adam optimizer and alternately optimizes the objective functions for the cross-domain transformation modules.

Approaching the Testing Phase

For any input query sketch and fashion photos in the retrieval dataset, the proposed cross-domain transformation module is used to transform them into the same domain.

Advanced Feature Extraction Techniques

Employing Symmetric CNN Architecture

Employs a symmetric CNN as the feature extraction module, utilizing VGG-16 pretrained on ImageNet as the backbone network.

The Process of Extracting Features

Extracts deep feature vectors for the sketch-based fashion photo and sketch retrieval streams, storing them in the database for further processing.

Measuring Cross-Domain Similarities

Evaluating Feature Similarity

Measures the similarity of the obtained feature vectors for both sketch-based fashion photo and sketch retrieval streams.

Resolving the Final Similarity

The relevant fashion photos from the dataset can be returned based on the final similarity measurement.

Experimentation and Performance Analysis

The experiments involve transformation and retrieval tasks for clothes, pants, skirts, and shoes.

The system demonstrates high retrieval accuracy and is evaluated through retrieval accuracy metrics.

This detailed process outlines the intricate steps and methodologies involved in the cross-domain transformation and feature extraction for fashion image retrieval, along with the testing and experimental results.

Sketch-Based Fashion Image Retrieval

Comparative Effectiveness in Matches

Top-1 vs. Top-10

Top-1 correct match poses a higher challenge than Top-10 correct match

Model secures high rankings more frequently in the Top-1 matches

Clothes and Accessory Retrieval Assessment

Outfit Components Retrieval

Correct clothing match ranks in the Top-1 96.6% of the time

Consistency in Bottomwear Retrieval

Pants retrieval rates for Top-1 and Top-10 are 92.1% and 96.6%, respectively

Efficiency in Skirt Retrieval

Skirts are accurately retrieved in the Top-1 and Top-10 at 91.0% and 97.1% respectively

Accuracy in Footwear Retrieval

True-match shoe image retrieval accuracy for Top-1 and Top-10 stands at 90.5% and 97.8%, respectively

Visual Representation of Retrieval Results

Inspection of Retrieval Outcomes

The left illustrates query sketches, while the right evidences the Top-10 retrieved fashion photographs

True-match photos within the Top-10 are predominantly ranked at the Top-1 position

Influence of Training on Model Performance

Training Iterations Impact Analysis

The accuracy of retrieval is sensitive to the amount of training iterations

Data displayed in Table 3 elucidate the fluctuation of retrieval accuracy correlated with training iterations

Optimal results surfaced at 470,000 iterations, with Top-1 accuracies for clothing, pants, skirts, and shoes affirming the figures of 96.6%, 92.1%, 91.0%, and 90.5%, respectively

Resilience to Sketch Intricacies

Handling a Range of Sketch Styles

Model consistently excels regardless of the sketches' degree of complexity or simplicity

Benchmarking Against Standard Models

Baseline Comparison Trials

Benchmarked over three separate datasets: the Fashion Image dataset, QMUL-shoes, and QMUL-chairs

Model's procedural transformation of sketches and photos to a unified domain notably enhances accuracy

Superiority in Fashion Image Retrieval

Surpasses the Sketchy baseline by margins of 25.1% in Top-1 and 2.4% in Top-10 retrieval accuracies

Advancements in Footwear Dataset Retrieval

Demonstrates superior performance, trouncing Dense-HOG + rank-SVM by 6.4% in Top-1 retrieval accuracy on QMUL-Shoes

Progress in Chair Dataset Retrieval

Exhibits an increase in Top-1 retrieval accuracy over ISN Deep + rank-SVM by 2.1% on the QMUL-Chairs dataset

Probing Ablation Studies

Gauging Model's Capabilities

Executed on Fashion Image, QMUL-shoes, and QMUL-chairs datasets

Findings encapsulated in Table 5

Contributions and Progression Pathways

Bolstering Fashion Image Dataset Resources

Enrichment of the field by addition of the Fashion Image dataset consisting of 36,074 sketch-photo pairs

Innovation in Retrieval Methodology

Pursuit of a novel algorithm dedicated to sketch-based fashion image retrieval stands driven by cross-domain transformation techniques, thereby refining retrieval accuracy

Scope for Enhancement

Recognizable limitations in transforming sketches to their photogenic counterparts; upcoming endeavors will involve augmenting the pool of fashion images alongside honing retrieval precision

Supplementing Information and Ethical Assurance

Dataset Acquirement

Fashion Image dataset procurable through correspondence with the author

QMUL-shoes and QMUL-chairs accessible via a designated website

Ethical Transparency

The contributing authors assert the absence of any conflict of interest

This assimilates the key observations and structured insights drawn from the source material, ensuring comprehensive retention of core elements and specifics.