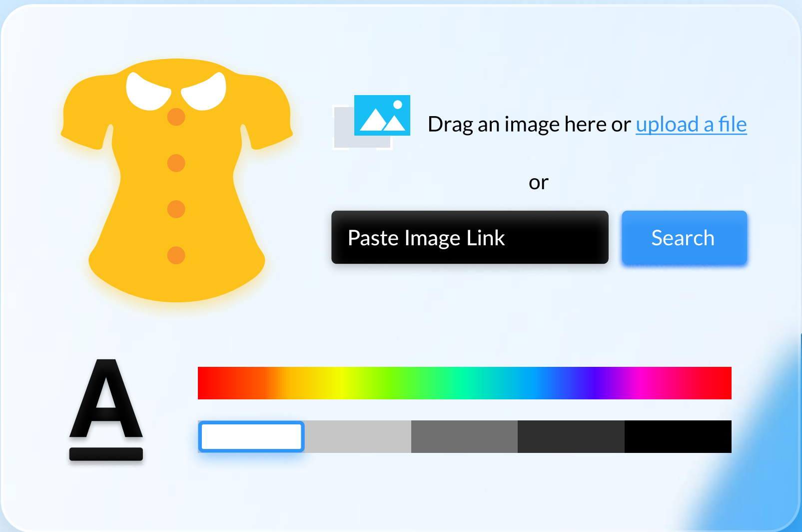
Netcore UNBXD

Netcore Unbxd's Visual Search Solution

Revealing the Future of Online Shopping



unbxd.com

Introduction to

Netcore Unbxd Visual Search

In 2013, 500B photos were taken, and by 2023, that number tripled to over 1.5 Trillion. By 2030, it is expected to hit 2 Trillion! 1.70 Trillion images per year means 54,400 images every second.

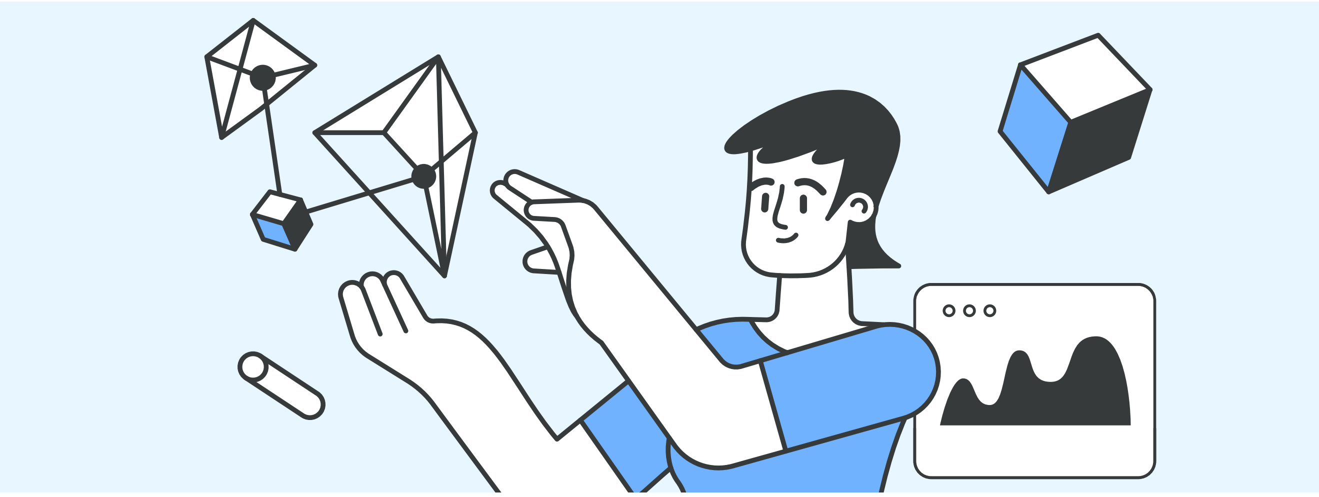
Why are ecommerce websites not leveraging the growth of photographs to scale their business? Imagine how easy it would be for your shoppers to hit the "buy" button if they can find what they want by easily and quickly uploading a picture.

Netcore Unbxd Visual Search uses AI and machine learning algorithms for feature detection and pattern recognition. Visual search is the end-to-end customer-facing product discovery experience. It is an essential tool to enhance the shopper's experience, ride a growing trend, and generate additional revenue.

Let us dive into the future of ecommerce and understand visual search's working principle, architecture, and benefits in this whitepaper. And how enterprises can use it to increase search relevance.

Visual Search: How it works

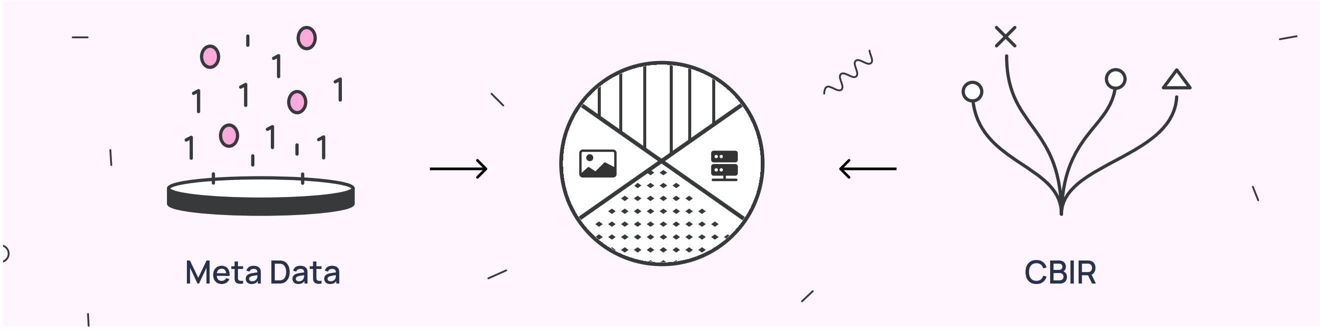
Simply put, Visual Search analyzes every aspect of the image, identifies product attributes, categorizes them, and then looks for similar products with as many identified attributes as possible. The more common attributes found, the more exact it is to the image it is comparing it to.



Visual search works on top of **vector search**: a shorter distance between vectors of any two items on the vector space means they would possess similar characteristics (distance calculated using algorithms like Cosine Similarity and Euclidean distance). But before we reach the phase of vectorization, the search bar needs to understand the images uploaded and interpret them.

There is a semantic gap between a digital footprint of an image and a human interpretation of the same image. This disparity tends to limit the accuracy and success of any search. Search engines need help understanding the context, meaning, and object within an image solely based on pixel information (information saved in the pixel of any picture).

This can be done using two techniques.



Using the Image Meta Data:

Image metadata is text information that includes color, shapes, and an array of other specifications.

All the images in the catalog/database are analyzed, and their metadata is kept in hand for reference when another image comes into the system. The newly captured metadata is then compared with the metadata of photos in the database/catalog. Search results are then ranked based on their similarity. Unbxd has optimized this process and made it remarkably accurate and fast. It takes Netcore Unbxd less than 4 seconds to scan a plethora of data and provide the best result.

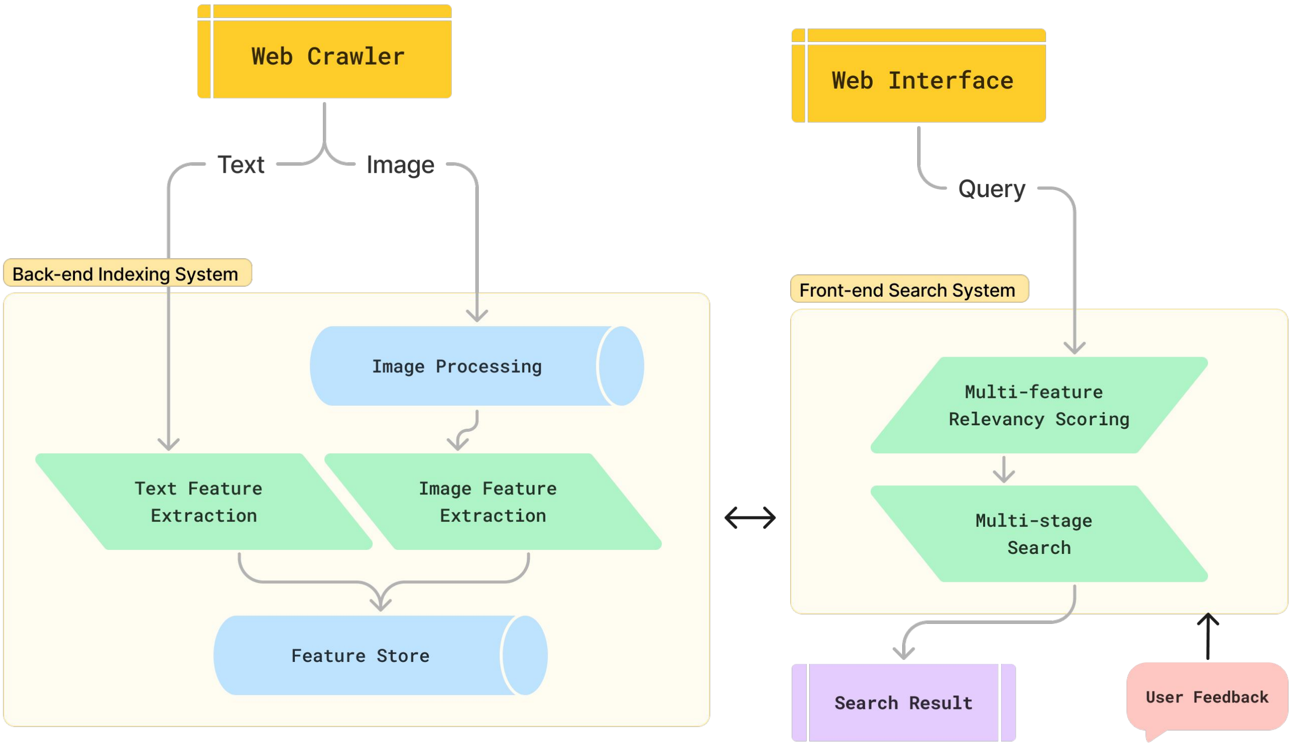
Content-Based Image Retrieval (CBIR):

While image metadata provides a decent amount of information, it's also essential for the search system to read the image in detail and identify the less-seen features.

This is where CBIR comes in. CBIR considers low-level visual features like **color**, **shape**, **texture**, and **spatial layout** to find images with similar contents. This process of uniquely identifying the characteristics of an object in an image is called fingerprinting.

Understanding the architecture

The back-end store collects product images from one or more sources to create a comprehensive database. Machine learning and AI algorithms process these images at the back end. The algorithm separates the background from the object through image segmentation; then, image features are extracted into text and image features using descriptors.



Identifying similarities and differences between images is essential for grouping objects through clustering. However, to accurately identify objects, assigning meaning to them is necessary. Photos must be explicitly named and labeled with semantic categories, which is crucial for image segmentation or classification. This process is essential for image comprehension and matching, which is the primary focus of visual search.

With Netcore Unbxd's generative AI algorithm, image segmentation or classification becomes hassle-free, and you will never have to worry about missing attribute names and descriptions.

Extracting data:

Images might look different, but if you scale them, they are just collections of pixels, some of which are illuminated or not illuminated. Visual search feeds the grid/matrix of an image into a neural network trained to interpret the input. Neural networks are effective machine-learning tools for analyzing and extracting visual information from images.

The first step would be deciding on a set of characteristics to measure the relevancy of the query image. The collection of attributes describing an appearance in the database is called descriptors. They fall into three major categories.

1. **Shape Descriptors**:
   a. Edge Histogram Descriptor (EHD) - The local-edge histogram is generated for each of the 4x4 blocks in the image, allowing EHD to capture both the object shape (external edges) and texture (internal edges) as intended.
   b. Region Shape Descriptor (RSD) - The bi-level foreground object mask undergoes a Complex 2D Angular Radial Transformation (ART) and extracts the resulting normalized coefficient magnitudes.

2. **Color Descriptors**:
   a. Color Structure Descriptor (CSD)- This method considers the color distribution and the local spatial structure of an image's colors. It constructs a histogram of quantized colors by analyzing how those colors appear in pre-defined structural elements.
   b. Color Layout Descriptor (CLD)- It captures the spatial distribution of the colors in an image. : The image is first converted into 8x8 "icon," which is then transformed by 8x8 DCT(Discrete Cosine Transform) in each color channel.

3. **Texture Descriptors**: Homogeneous Texture Descriptor (HTD) describes the directionality, coarseness, and regularity of patterns in texture images.

Following the selection of features, the next step is to measure their performance in product image search. We select the **nearest neighbors (NN)** in the inventory based on individual features and then measure the precision-recall curve based on ground truth.

It will be interesting to note that texture(HTD) and shape descriptor (EHD/RSD) will recall better and more accurate results than compared to color (CSD and CLD).

Now we will deepen our understanding of AI technology and machine learning in visual search.

As we have learned, any image represents pixels; all images fed into the database get broken down into numbers called vectors. These vectors get passed through a neural network called CNN (Convolutional Neural Network).

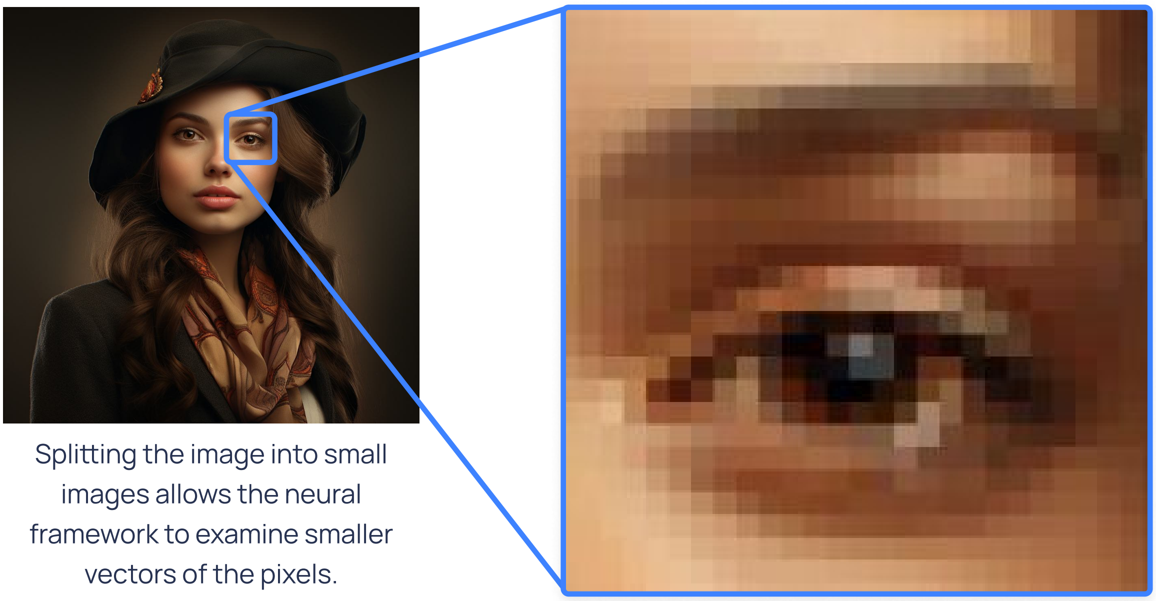
A **Convolutional Neural Network (CNN)** employs a model consisting of a statistical function sequence that continuously analyzes and updates the pixelated numerical vector until the machine can accurately recognize and classify the image.

During the learning operation, many images are entered into the grid, enabling the model to adapt, fine-tune itself, and experience subtle modifications of its internal numerical values known as **weights**. As the learning advances, the model slowly modifies its weights to strike a near-perfect balance, eventually enabling accurate recognition of future pictures and not compromising the win rate.

The transformed pixel (vectors) get represented in the form of numbers. The vector represents a specific range of pixels: a grayscale spectrum for black and white images or a three-dimensional RGB scale for colored photos. Each number (pixel) within the vector corresponds to a value within the grayscale or color scale.

The fundamental concept involves feeding the vector into the model and allowing its numbers to transform as they multiply through the network. The network comprises multiple layers, each interpreting the vector numbers as representations of distinct features. As the vector passes through each layer, new features add details to the image.

As mentioned earlier, the image transforms into a pixel (vector). However, rather than analyzing the picture, the machine breaks it into smaller vectors, typically in 5x5 mini-sections. For instance, a small 10x10-pixel image would be divided into 25 smaller images, each representing a distinct mini-section.



Splitting the image into small images allows the neural framework to examine smaller vectors of the pixels.

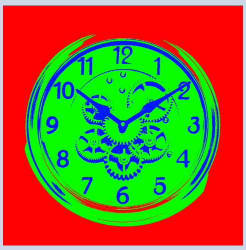
This image processing technique breaks the image into smaller vectors from the new vectors for better understanding.

The Input vector (a) gets transformed into the output vector (b) with continuous passing through the framework.

If you are starting to create your dataset or have an existing one, the most important contributing factor would be having a clean and clear dataset. Clarity has to do with the focus and lighting of the image, and cleanliness has to do with clutter or noise.

Clarity

Query Image



Object Map

A clear image - Note the details visible in the image itself, and the clean lines that define the shape of the object.

Query Image

Object Map

A not so clear Image - Note the lack of focus and that there is no real image in the object map to identify.

A clean image - White background, nothing but the image of interest.

A not so clean image - Lack of Visual focus. What are we looking for?

The benefits of visual search

1. **Increased accuracy of results:** Best-in-case visual search algorithms can provide accuracy of up to 95%. Advanced visual search tools can do more than just identify and interpret products in an image. They can also automatically recognize the category and type of a product, as well as the age and gender of models. Moreover, these tools can provide even more detailed information, including the physical attributes of items, such as their material, texture, shape, style, and occasion. This level of detail enables brands and retailers to deliver highly personalized search results that match precisely what a shopper seeks, including the tactile experience of an item's fabric.

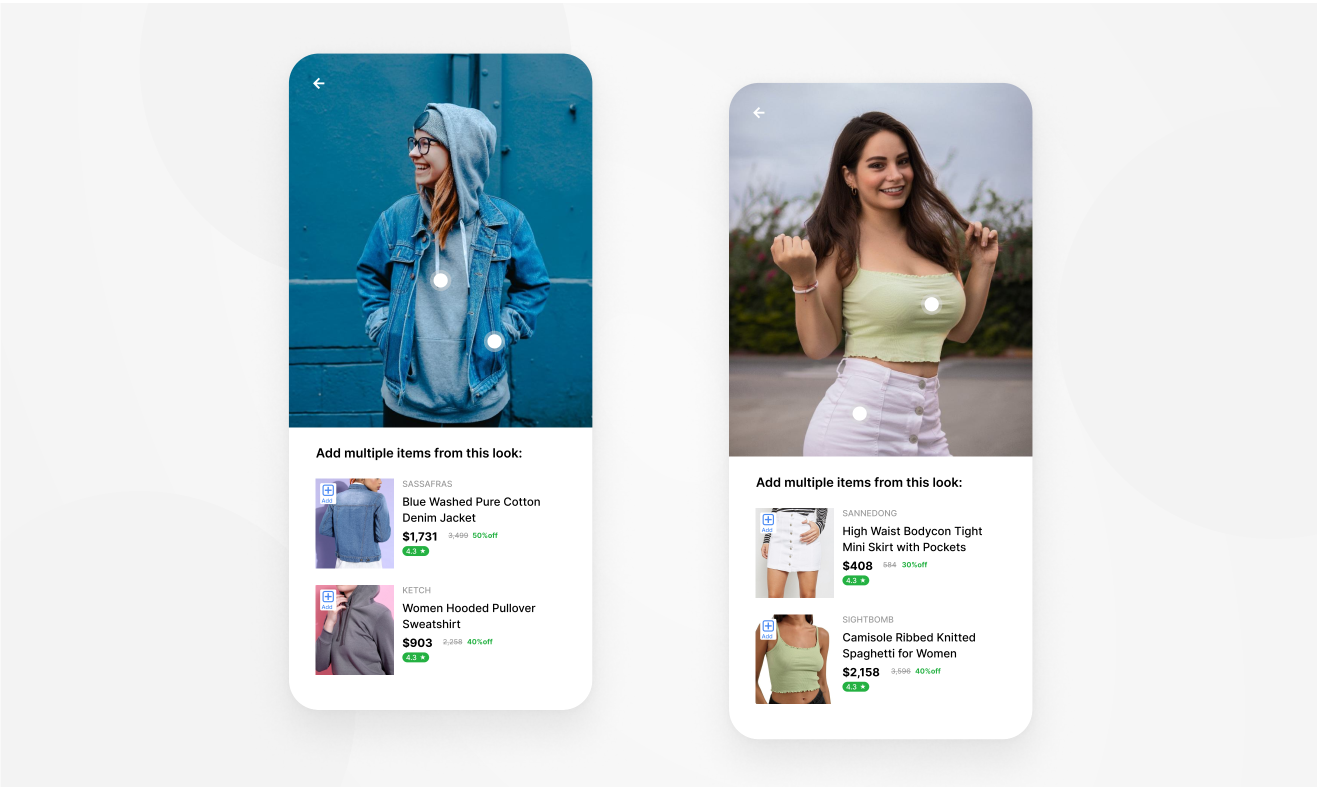
2. **Get an exceptionally similar product recommendation:** Visual search technology allows ecommerce sites to suggest products that a buyer may find appealing. Numerous product recommendation engines are available that use various data points to determine the most appropriate products to display, and computer vision is one of them. AI engines can quickly provide visitors with the most similar items they seek by analyzing visual cues such as color, style, shape, and other filters.

3. **Quick route from search to sell:** The easier it is for a customer to find the same or similar product on an ecommerce website, the less friction they will have before making a purchase.

4. **Shop the look:** Celebrity fashion and influencer trends have become popular sources of inspiration for Gen Z and millennials. As a result, retailers and brands are now utilizing visual technology to provide a more seamless and personalized buying experience for trend-conscious consumers.

   One such technology is "Shop the Look," which enables users to select multiple items within a picture that complement each other.

   For instance, imagine viewing a photo of a celebrity on Instagram, where you can click and purchase the t-shirt, skirt, shoes, and accessories they are wearing from the same image. The furniture industry uses the same principle; a single picture can recommend an entire catalog.

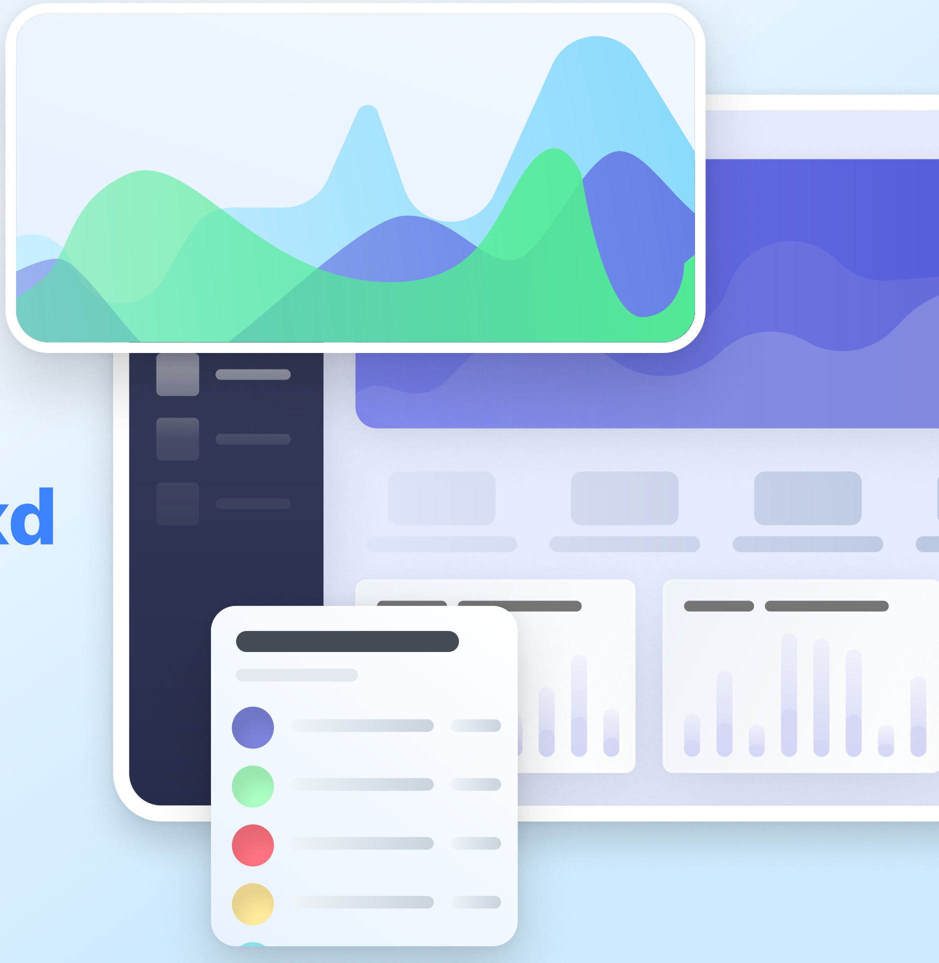


The Unbxd Innovation

Netcore Unbxd Visual search will accept images from anywhere and can extract data. Whether sourced from the web, their gallery, or captured impulsively, it accommodates multiple file formats, handles varying image qualities, and delivers visually similar results in just 3 seconds.

AI-powered computer vision algorithms consider attributes like **category**, **material**, **color**, and **occasion** and give the most accurate results. Unbxd's AI capabilities enhance the accuracy and relevance of the search results.

In conclusion, implementing visual search technology represents a significant advancement in product discovery. By harnessing the power of computer vision, image recognition, and deep learning algorithms, visual search has revolutionized how users search for and discover products. Through the ability to search with images instead of text, users can effortlessly find relevant products, explore similar options, and gain inspiration with just a few clicks. The accuracy, efficiency, and enhanced user experience provided by visual search contribute to increased customer satisfaction, improved conversion rates, and, ultimately, the success of businesses. As technology evolves and improves, visual search holds immense potential for reshaping the landscape of product discovery, making it an indispensable tool for modern ecommerce platforms and beyond.



Visual Search

Netcore Unbxd White Paper 2024

Netcore UNBXD

1710 S. Amphlett Blvd
Suite 124 San Mateo,
CA 94402

sales@unbxd.com
support@unbxd.com

+1 (650) 282-5788