
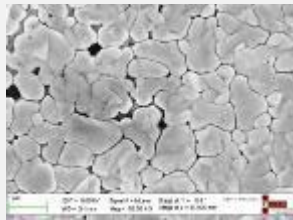


VQA T1 Evaluation

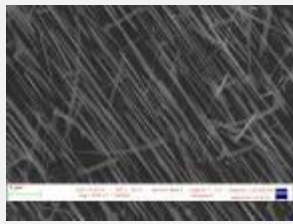
Performance of ChatGPT, LLaVA, and Gemini (left to right) on VQA tasks of describing an SEM image from each class in the NFFA dataset. Correct key responses are colored in green, wrong responses in red, and mixed or uncertain in orange. Extraneous information has been replaced with ellipses in some places. The ground truth about the image, and the text from the annotation was read and, was written by our domain expert. The responses have been scored by our domain expert on a scale of 0 to 5 and have been presented as three numbers for the three models respectively in the Scores column.

ChatGPT	LLaVA	Gemini	Scores
	L0_1c34778dec5a1701d689a7067c4813eb.jpg Class: Porous Sponge		
	Ground truth: The SEM image depicts a porous material with irregularly shaped holes hundreds of microns wide, within a material consisting of spherical particles around 10 μm in diameter. Scale bar = 100 μm , Magnification = 250 X, WD = NaN, EHT = 3.00 kV, Stage Z = 48.000 mm, Logo = Zeiss.		
<p>The SEM image shows a porous material with large, irregularly shaped holes distributed throughout the surface. The edges of the holes are lined with smaller spherical particles, giving a rough texture to the surface. The scale bar indicates features in the micrometer range, with a magnification of 250x. The image was captured using a Secondary Electron (SE2) detector at an electron high tension (EHT) of 3.00 kV. The image was taken on July 18, 2014, at 16:13:06. The porous structure suggests applications in filtration or catalysis. The image was taken using a Zeiss microscope.</p>	<p>The image is a black and white scanning electron microscope (SEM) micrograph. It shows a close-up view of a surface with numerous irregularly shaped particles or structures. These particles have a rough, uneven texture and are densely packed together, creating a pattern that resembles a natural mineral or a composite material. The image is taken at a magnification of 250x, and the scale bar at the bottom indicates 100 μm, suggesting that the particles are on the micrometer scale.</p> <p>The SEM settings are also provided at the bottom of the image, including the voltage (EHT) at 3.00 kV, the signal (Signal A) at SE2, the stage tilt angle (Stage T) at 0°, and the brightness at 34%. The contrast is set to 66.6%, and the aperture size is 30.0 μm. The stage is at a height of 48,000 μm. The image is labeled with the Zeiss logo, indicating the brand of the microscope used to capture the image.</p>	<p>The image you sent is a black and white photo of a surface with many holes in it. The holes are all different sizes and shapes, and their edges look rough and uneven. Some of the holes appear almost circular, while others are long and jagged. The material surrounding the holes also has a bumpy, uneven texture.</p> <p>The text along the bottom of the image indicates that it was captured with a scanning electron microscope (SEM) at a magnification of 250x. The date the image was taken was July 18, 2014. Other information about the microscope settings and the sample itself is also included in the text, but it is too technical to understand without specific knowledge of SEMs.</p> <p>Given the variety of shapes and sizes of the holes, it's difficult to say what this material is or how the holes were created. It is possible that this is a natural material, like rock or coral, that has been eroded or weathered over time. Alternatively, it could be a man-made material that has been subjected to some kind of corrosive or abrasive process.</p>	5 1 2
	L1_0f99a90d314c84b87014c042a746c276.jpg Class: Patterned Surface		
	Ground truth: The SEM image shows a microfabrication of something that might be used for calibrations, with finely resolved patterns in the form of curves, parallel lines, and square arrays. Given by the size of the lines, advanced lithography or nanolithography techniques must have been used in its production. Scale bar = 10 μm , Magnification = 2760 X, WD = 4 mm, EHT = 5.00 kV, Stage Z = 48.482 mm, Logo = Zeiss.		
<p>The SEM image displays a microfabricated calibration grid with various patterned regions. These patterns include square arrays of dots, linear features, and circular arcs. The structures are highly regular and precise, indicating the use of advanced lithography techniques. The scale bar shows features in the micrometer range, with a magnification of 2,760x. The image was captured using a Secondary Electron (SE2) detector at an electron high tension (EHT) of 5.00 kV. The working distance (WD) is 4 mm. The grid is likely used for calibration or testing in microscopy and nanofabrication. The image was taken using a Zeiss microscope.</p>	<p>The image shows a close-up of a surface with a pattern of circular indentations. The surface appears to be a metallic or composite material with a matte finish. There are three distinct areas of circular indentations, each with a different density and size of the circles. The circles are arranged in a grid-like pattern, and the indentations are uniform and consistent within each area.</p> <p>Below the patterned surface, there is a white label with text and measurements. The text includes "10 μm," "EHT = 5.00 kV," "WD = 4 mm," "Signal A - SE2," "Stage at T = 0°," "Stage at T = 482°," and "Aperture Size = 30.0 μm." The label also features a logo at the bottom right corner, which is a stylized letter 'Z' inside a circle,</p>	<p>The image you sent is a close-up of a grid of squares and circles arranged in a repeating pattern on a black background. The squares and circles are all roughly the same size and appear light gray. The text along the bottom of the image specifies it was captured with a scanning electron microscope (SEM) and provides details about the magnification and other settings used to create the image.</p> <p>SEM images are used to image the surface of a sample by scanning it with a focused beam of electrons. The way the electrons interact with the atoms in the sample creates signals that provide information about the surface's topography and composition.</p> <p>Based on the presence of the grid, it's possible this image is a micrograph of a material being</p>	5 1 3

ChatGPT	LLaVA	Gemini	Scores
	<p>suggesting it might be a brand or company logo.</p> <p>The style of the image is technical and appears to be a microscopic or scanning electron microscope (SEM) image, as indicated by the scale and the reference to "Signal A - SE2," which is likely a scanning electron microscope signal. The image is likely used for scientific or technical analysis, possibly in materials science or engineering.</p>	<p>used in the development of microelectronics or microsensors. The repeating pattern of the squares and circles could be a test structure used to calibrate the SEM or to assess the quality of the material during its development.</p>	
	<p>L2_0ac309ffe6244484d9d33d119bc686a2.jpg</p> <p>Class: Particles</p> <p>Ground truth: The SEM image shows uniformly sized, spherical and scattered nanoparticles with some level of clustering or agglomeration of a few particles. Some isolated particles can also be seen.</p> <p>Scale bar = 200 nm, Magnification = 234190 X, WD = 3 mm, EHT = 10.00 kV, Stage Z = 48.920 mm, Logo = Zeiss.</p>		
<p>The SEM image shows a distribution of small, spherical particles dispersed across the field of view. These particles are uniformly sized and appear to be evenly spread out, with some clustering in small groups. The scale bar indicates features in the nanometer range, with a magnification of 234,190x. The image was captured using an InLens detector at an electron high tension (EHT) of 10.00 kV. The working distance (WD) is 3 mm. The particles are likely nanoparticles, and their uniform size and distribution suggest potential applications in nanotechnology or materials science. The image was taken using a Zeiss microscope.</p>	<p>The image is a microscopic view of a sample, likely biological in nature, captured using a scanning electron microscope (SEM). The sample appears to be a collection of small, spherical particles or cells, scattered across the surface of the sample. The particles are densely packed, with some areas having a higher concentration than others. The image is in grayscale, with the particles standing out against the darker background.</p> <p>At the bottom of the image, there is a label with technical information about the image. It includes the magnification level (249.19 kX), the sample size (3 mm), the signal (A-In Lens), the stage tilt (0°), the stage position (482.90 mm), the scan speed (9), and the aperture size (30.00 µm). The label also mentions "EHT = 10.00 kV," which likely refers to the accelerating voltage used in the SEM. The logo of Zeiss, a company known for manufacturing optical and optoelectronic equipment, is visible at the bottom right corner of the label.</p>	<p>The image you sent is a close-up of a black background with many small, white dots scattered across it. The even distribution of the dots suggests a regular pattern, rather than a random scattering. The text along the bottom of the image likely provides information about the creation of the image, but it is technical and difficult to understand without specific knowledge of scanning electron microscopes (SEM).</p> <p>Given the regularity of the dots, it's possible this image is a micrograph of a surface with a patterned texture, like a sensor or a filter. The even spacing of the dots could represent tiny holes or bumps on the surface. Without more information it's difficult to say for sure.</p>	5 2 1
	<p>L3_77ec686d269e87e1cb3eb05c890a4cc3.jpg</p> <p>Class: Coated Films</p> <p>Ground truth: The SEM image shows closely packed grains geometrically shaped with rounded edges and clear boundaries, size of the order of a micron, possibly indicating a crystal structure on a flat surface. It is difficult to pinpoint that this is a coated film.</p> <p>Scale bar = 1 µm, Magnification = 50000 X, WD = 3.1 mm, EHT = 5.00 kV, Stage Z = 47.000 mm, Logo = IOM.</p>		
<p>The SEM image shows a microstructure with polygonal grains, characteristic of a crystalline material. The grains appear closely packed with distinct grain boundaries. The scale bar indicates that the features are in the micrometer range, with the magnification set at 50,000x. The image was captured using an InLens detector at an electron high tension (EHT) of 5.00 kV. The working distance (WD) is 3.1 mm. The date of capture is February 3, 2016, at the Istituto Officina dei Materiali. The grains exhibit smooth surfaces and various sizes, suggesting polycrystalline nature.</p>	<p>The image is a microscopic view of a sample, likely a biological or material specimen, captured using scanning electron microscopy (SEM). The sample appears to be composed of numerous irregularly shaped, white, and translucent particles that are densely packed together, forming a pattern that resembles a cracked or shattered surface. The particles vary in size and shape, with some appearing more rounded while others are more irregular.</p> <p>At the bottom of the image, there is a scale bar indicating 1 µm (micrometer), which provides a reference for the size of the particles. The image also includes a set of technical parameters and data:</p> <p>EHT (Electron High Tension) = 5.00 kV (kilovolts)</p> <p>Signal A = In-Lens</p>	<p>The image you sent me is actually a black and white scanning electron microscope (SEM) image of a rock, not a piece of metal. SEM images are used to image the surface of a sample by scanning it with a focused beam of electrons. The electrons interact with the atoms in the sample, producing various signals that contain information about the sample's surface topography and composition.</p> <p>The text in the image indicates that the image was taken at 5.00 kV and a magnification of 50,000 times. The date the image was taken was February 3, 2016. The image was captured by a machine from Giom, Istituto Officina dei Materiali.</p> <p>I can't tell what type of rock it is from this image, but I can tell you that it is a very rough</p>	4 3 0

ChatGPT	LLaVA	Gemini	Scores
	<p>Stage at T = 0.0° Stage at Z = 47,000 mm Date: 3 Feb 2016 The image is labeled with "iOM" and "Instituto Officina," which could be the name of the institution or the software used to analyze the SEM data. The overall style of the image is scientific and technical, with a focus on providing detailed information about the sample's structure and composition.</p>	<p>and uneven surface. There are all sorts of bumps and crevices on the surface of the rock.</p>	
	<p>L4_1edc8f0dca1bc85dcb3e7ed28654a017.jpg Class: Powder</p> <p>Ground truth: The SEM image shows a dense agglomeration of spherical particles, which could be part of a porous surface. The material is composed of nano-sized particles coming together to form irregular surface. The surface is highly rough in the nm range and shows a lot of 3D depth in particle packing. Scale bar = 200 nm, Magnification = 100000 X, WD = 2.4 mm, EHT = 3.00 kV, Stage Z = 47.897 mm, Logo = TASC.</p>		
<p>The SEM image shows a highly textured surface composed of closely packed, irregularly shaped particles. These particles appear to form a porous, granular structure. The scale bar indicates features in the nanometer range, with a magnification of 100,000x. The image was captured using an InLens detector at an electron high tension (EHT) of 3.00 kV. The working distance (WD) is 2.4 mm. The image was taken on December 6, 2013, at 15:10:26. The granular texture suggests high surface area, which could be useful for applications in catalysis or adsorption. The image was taken at the TASC laboratory using a Zeiss microscope.</p>	<p>The image is a grayscale scanning electron microscope (SEM) image of a material with a complex, densely packed structure. The material appears to be composed of numerous small, irregularly shaped particles or aggregates that are tightly packed together, creating a textured, almost crystalline appearance. The particles vary in size and shape, with some appearing more spherical while others are more irregular. At the bottom of the image, there is a label with various parameters and settings that were used to acquire the SEM image. These include the date and time the image was taken, the accelerating voltage (EHT), the working distance (WD), the brightness, the stage tilt angle, the stage position, the magnification, the signal type, the lens contrast, the tilt compensation, and the aperture size. The label also includes the logo of the company that likely manufactured the SEM, which is TESCAN. The image is a scientific representation, typically used in materials science and engineering to study the microstructure of various substances. It provides detailed information about the morphology and composition of the material, which can be crucial for understanding its properties and applications.</p>	<p>The image you sent is a black and white photomicrograph, likely taken with a scanning electron microscope (SEM), which shows a cluster of biological cells. The cells are arranged in a circular pattern and vary in size and shape. Some cells are round and smooth, while others are more elongated and wrinkled. The text along the bottom of the image contains information about the settings used to capture the image, including the magnification (100,000x), the date (December 6, 2013), and other technical details about the microscope settings. Given the presence of biological cells, it is possible this image is a micrograph of a tissue sample. However, without additional context, it is difficult to say for sure what type of cells these are or what tissue they come from.</p>	5 3 0
	<p>L5_0b88b7fc72ff01ff3802cdc4edf5d6c4.jpg Class: Tips</p> <p>Ground truth: The SEM image displays pointed tip or a tapering needle like structure with tip size far below 1 μm, exhibiting surface texture likely originating from another material stuck to the needle. The fine detail and sharpness suggest uses in precision tools or probes. Scale bar = 2 μm, Magnification = 10000 X, WD = 7.6 mm, EHT = 5.00 kV, Stage Z = 49.000 mm, Logo = Zeiss.</p>		
<p>The SEM image shows a sharp, needle-like structure with a tapered end. The surface of the needle exhibits some roughness and irregularities. The scale bar indicates features in the micrometer range, with a magnification of 10,000x. The image was captured using a Secondary Electron (SE2) detector at an electron high tension (EHT) of 5.00 kV. The working distance (WD) is 7.6 mm. The stage is tilted at 45 degrees. The image highlights the detailed surface morphology and the</p>	<p>The image is a microscopic view of a needle-like object, captured using a scanning electron microscope (SEM). The needle appears to be made of a material with a rough surface texture, and it is positioned diagonally across the frame. The SEM has been used to provide a detailed view of the needle's surface, revealing its microscopic structure. The image is in black and white, emphasizing the contrast and detail of the needle's surface. The SEM has been set to a magnification of 10.00 kX, allowing for a</p>	<p>The image you sent is a close-up microscope image of a needle, likely taken with a scanning electron microscope (SEM) due to the high magnification and level of detail. The needle appears very thin and sharp, with a smooth, tapered tip. The text along the bottom of the image indicates it was captured at a magnification of 10,000 times (10.00 KX) and a working distance of 7.6 millimeters (WD=7.6~mm). There are also other technical details listed such as the accelerating voltage</p>	5 4 4

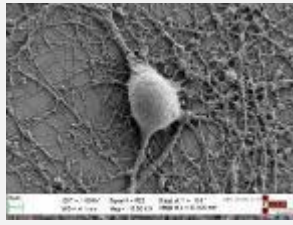
ChatGPT	LLaVA	Gemini	Scores
sharpness of the tip, which could be used for applications requiring fine probes or sharp tools. The image was taken using a Zeiss microscope.	close-up view of the needle. The needle is 2 μm in length and has a width of 7.8 μm . The SEM has been set to a stage position of 45.00 degrees, and the needle is located at a stage position of 48.00 μm . The SEM has an aperture size of 30.00 μm . The image provides a detailed view of the needle's surface, revealing its microscopic structure and texture.	(EHT 5.00 KV) and the aperture size (Aperture Size = 30.00 μm). The needle itself is difficult to identify from this image alone. It could be a sewing needle, a hypodermic needle, or some other type of needle entirely.	



L6_00d0f158ae4f1993980e7ee8375a231f.jpg
Class: Nanowires

Ground truth: The SEM image depicts highly aligned, rigid rod structures with lengths more than ten times the diameter, likely nanowires or fibers inside a polymer matrix, aligned in a parallel arrangement with some irregularities in the material shape and orientation. The uniform distribution and high aspect ratios suggest potential applications in nanotechnology or materials science.
Scale bar = 1 μm , Magnification = 40000 X, WD = 4 mm, EHT = 10.00 kV, Stage Z = 47.957 mm, Logo = Zeiss.

<p>The SEM image shows a dense array of elongated, needle-like structures, possibly nanowires or fibers, arranged in a somewhat parallel manner. These structures vary in length and appear to intersect at various points. The scale bar indicates that the features are in the micrometer range, with a magnification set at 40,000x. The image was captured using an InLens detector at an electron high tension (EHT) of 10.00 kV. The working distance (WD) is 4 mm. The structures appear to be uniformly distributed and exhibit high aspect ratios, indicating potential applications in nanotechnology or materials science. The image was taken using a Zeiss microscope.</p>	<p>The image is a grayscale scanning electron microscope (SEM) image of a sample. The sample appears to be composed of thin, elongated structures that resemble fibers or strands, which are densely packed and oriented in various directions. The structures have a smooth, possibly crystalline surface, and they are set against a dark background, which is typical for SEM images. At the bottom of the image, there is a label with technical information about the image. It includes the magnification level (40.00 X), the signal used (A-In Lens), the stage tilt angle (0.0°), the stage position (Z=47.58 mm), the scan speed (9), and the aperture size (30.00 μm). The label also mentions the voltage (EHT = 10.00 kV) and the working distance (WD = 4.0 mm). The logo of Zeiss, a company known for producing optical and optoelectronic equipment, is visible at the bottom right corner of the label. The overall style of the image is scientific and technical, with a focus on the microscopic details of the sample.</p>	<p>The image you sent is a microscope image of a metal surface, likely magnified thousands of times. The surface is filled with irregular lines and shapes that vary in size. Some of the lines are thin and smooth, while others are thicker and more jagged. There are also some lighter colored areas throughout the image, which may be due to different materials present on the surface or the way the light interacted with the sample during imaging. The text along the bottom of the image indicates it was captured with a scanning electron microscope (SEM) at a magnification of 40,000 times. The presence of different features and textures suggests the surface may have been manipulated or treated in some way.</p>	541
---	--	---	-----



L7_97b73f055d76fb834374f025d3a83802.jpg
Class: Biological

Ground truth: An SEM image of a neuron showing the cell body (called soma) in the center with thin filamentous structures (called neurites) extending out from it.
Scale bar = 2 μm , Magnification = 10000 X, WD = 4.1 mm, EHT = 1.00 kV, Stage Z = 47.000 mm, Logo = IOM.

<p>The image you've provided appears to be a scanning electron microscope (SEM) image... In this image, we see a large, rounded structure that could be a cell or a particle, surrounded by an intricate network of fibrous or filamentous structures. The background suggests a mesh-like or highly textured material, possibly extracellular matrix or some kind of synthetic scaffold. The rounded structure is significantly larger than the surrounding meshwork, highlighting it as the central figure in this image. The labels on the image, like "EHT = 1.00 kV" (electron high tension or the accelerating voltage), "WD = 4.1 mm" (working distance), and "Mag = 10.00 KX" (magnification level), provide technical information about how the SEM image was captured. The date stamped is "29 Dec 2015", indicating when the image was taken.</p>	<p>The image appears to be a microscopic view of a biological structure, possibly a cell or a tissue sample. It is a grayscale image with a high level of detail, showing a network of thin, thread-like structures that could be fibrous proteins or other components of the cell. At the center of the image, there is a spherical object that stands out from the surrounding network. This object has a smooth, rounded surface and is connected to the network by what looks like a thin stalk or membrane. The image includes a label... The magnification is specified as 10,000x... The scale bar... 2 μm (micrometers) corresponds to the length of the bar. Additional information includes the signal strength (Signal A = SE2), the stage position at T = 0°, and the stage position at Z = 47,000 mm. The date of the image is also provided, which is</p>	<p>The image displays a scanning electron microscope (SEM) picture of a neuron, likely cultured in a laboratory environment. The neuron is characterized by a sizeable cell body, also known as the soma, from which numerous thin, elongated structures called neurites extend.</p>	445
--	---	--	-----

ChatGPT	LLaVA	Gemini	Scores
	December 29, 2015. The logo of the Instituto Officina dei Materiali...		



L8_1a7066d761da10e1366cddcc354d7a7f.jpg
Class: MEMS

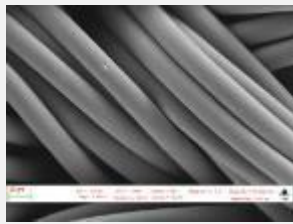
Ground truth: The SEM image shows a micro circuit showing rectangular conductive materials interconnected by thin curved conducting lines. The rectangular regions seem to be around 20 μm wide. The well-defined edges indicate precise fabrication techniques used in its production.
 Scale bar = 20 μm , Magnification = 1000 X, WD = 10.7 mm, EHT = 5.00 kV, Stage Z = 45.000 mm, Logo = TASC.

The SEM image depicts a patterned microelectronic structure. The image shows interconnecting lines and pads likely made of a conductive material, arranged in a symmetrical design typical of electronic circuits. The scale bar indicates features in the micrometer range, with a magnification of 1,000x. The image was captured using a Secondary Electron (SE2) detector at an electron high tension (EHT) of 5.00 kV. The working distance (WD) is 10.7 mm. The capture date is May 29, 2013, at 9:34:49. The structure appears well-defined with sharp edges, indicating high precision in fabrication.

The image is a grayscale photograph of a microscopic view of a circuit board. The circuit board has a complex pattern of lines and shapes, including rectangles and triangles, which are typical of electronic components and pathways. The background is dark, providing a stark contrast to the lighter lines and shapes on the circuit board. At the bottom of the image, there is a white strip with text that provides information about the photograph. The text includes the date "29 May 2013," the magnification "1.00 KX," the signal "A," the stage "T = 0°," the stage at "Z = 45,000 mm," the time "9:34:49," the brightness "48.0%," the contrast "35.0%," the tilt angle "0°," the tilt correction "Off," the aperture size "30.0 μm ," and the TASC logo. The logo suggests that the image may have been taken with a TASC microscope or software. The text and logo are in a sans-serif font, and the overall style of the image is technical and informative.

The image you sent is a black and white close-up of a circuit board. The circuit board is a flat, rigid plate that contains electronic components. These components are connected together by thin wires, or traces, that make up the circuits. In the image, the circuit board appears light gray against a black background. The circuit board is filled with a complex network of winding traces. The traces branch out and connect to various rectangular shaped components on the surface of the circuit board. The text in the bottom right corner of the image provides information about the conditions under which the image was captured, but it is too technical to understand in detail.

5 4 4



L9_67328b7fc592b7625489dc1e9c12ba22.jpg
Class: Fibers

Ground truth: An SEM image of synthetic fibers showing bundles of smoothly surfaced fibers in an interwoven formation. Scale bar = 20 μm , Magnification = 1430 X, WD = 4 mm, EHT = 1.00 kV, Stage Z = 47.509 mm, Logo = TASC.

The structures could represent fibers, tubes, or filaments of some material, whether biological (like plant fibers or extracellular matrices) or synthetic (such as polymer fibers or nanotubes). Scale bar "20 μm " (micrometers), indicating the scale for size comparison. Additionally, the image data provides technical details: "EHT = 1.00 kV" (electron high tension or the accelerating voltage), "WD = 4 mm" (working distance), "Mag = 1.43 KX" (magnification)...

The image... possibly a fiber or a bundle of fibers, captured using scanning electron microscopy (SEM)... are tightly packed together, creating a dense, layered pattern... label with various technical specifications... It includes the magnification (1.43kX), the signal (SE2), the stage position at T (0.0°), the stage position at T (45.69°), and the aperture size (30.00 μm) ... voltage (1.00 kV), the working distance (4.0 mm), the signal brightness (60.0%), and the contrast (43.2%). The label also mentions "TASC," which could be an acronym for a specific model or feature of the SEM equipment used to capture the image....

The image displays a microscopic view of interwoven fibers, likely captured using a scanning electron microscope (SEM) due to the detailed surface texture and grayscale image. The fibers have a cylindrical structure and a relatively smooth surface with some visible irregularities. The arrangement of the fibers suggests a woven or braided material... The specific type of fiber and its application are not evident from the image alone.

3 4 5