# **Prompt Exploration**

# **MODÈLES**

Les modèles testés sont les suivants :

- llava:7b	(7b)
- llava:13b	(13b)
- llava:34b	(34b)
- llava-llama3	(8b)
- llama3.2-vision	(11b)
- gemma3:4b	(4b)
- gemma3:12b	(12b)
- gemma3:27b	(27b)

Les résultats observés sur 13 prompts sont les suivants :

- llama3.2-vision a obtenu les meilleurs scores pour chaque prompt testé
- les modèles llava ont tendances à obtenir les pires résultats en moyenne

gemma3:12b	gemma3:27b	gemma3:4b	llama3.2-vision	llaya	llava-llama3	llaya:13b	Ilaya:34b	
31	30	29			22	28	20	
26	27	18	33	19	17	21	14	
28								
26	33	21	30	25	24	25	17	
29	33	23	39	25	26	26	22	
18	11	13	23	12	11	11	11	
22	22	23	32	17	22	20	23	
27	32	24	37	24	13	25	22	
29	31	. 27	36	27	22	25	26	
34	35	28	38	17	26	23	21	
36	35	27	36	24	24	26	25	
30			37	26				
31			33					
367	381	306	449	288	278	306	274	TOTAL
3	2	4	1	6	7	4	8	
Max	Max-1	Max-2	Max-3-5		Min+3-5	Min+2	Min+1	Min

- la taille des modèles semble avoir un impact assez imprévisible, par exemple pour llava on a eu les résultats suivants llava:34b < llava:7b < llava:13b, donc le plus grand modèle a eu les pires résultats et le modèle de taille moyenne a eu les meilleurs résultats. Pour gemma3 les scores ont suivi la taille du modèle.

Mais 1	les tai	gemma3:27b semma	gemma3:4b na	ıt réu	ssir/ra	iter su	ır diff	érent	es in	nages,	gemma3:12bd	gemma3:27bax	gemma3:4b					
***	generich	generality	general to	Named 2 street	-	Acceptant					parent.	position		Name of Street	-			
Bell in a colorino agricoro visione del dispublica y a per la colorida del colorida del colorida del colorida del colorida del colorida del colorida del colorida del periodi del colorida																1	-	
Mr 4: 7 children approprie children for described and in terminal prior of	-	-	hos	tu.	-	Na	ha	-		At a c white agreem wheth the c c con a single star of the part page.	-	No.	-	-			No.	-
Mountain state and late, and late and l	England 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1															-	-	
Beautiful and shall you to be to be a few of the control of the co																-		
Should, the case is the discharge																~		
Section Control of the Control of th															~			
What's the appet to the take magnifestate one with simple soul															-			
Water I has appear to the the company of part files. On the part of part files of contract bearings of contract appear contract.	No	ha .	-	2	-	No.	ter	No.		What is the Japan in the state of the stage	-	~	-	-	~	~	-	~
We are a varieties flament are the restricting are they derive the company of the company of the company of the company of the company of the company of the company of the company of the company of the	No	No.	-	No.	No.	No.	tu .	No.		We are a round found on the order per per per per per per per per per p	~	1	-	1	-	-	-	-
No se excisión based en para el calde que de cardo el calde que de cardo el calde que de cardo el calde que de cardo el calde que el calde que el calde el ca	No		No.	to .	he .	No.	to	No.		We see which family on a con- part of the party of the call, and the party of the call, and the call of the call, and the call of the call	~	1	-	1	-	-		-
No an acceptant famous or per or other gap day, can or for a per or other acceptants. The formation of the contract of contracting famous acceptants.	No.	No.	-	2	-	-	-	he		No per a related formed any per per per per per per per per per per	-	1	-	-	-	-	-	-
We as a visite faunt or some other gar the principal or agent of the principal or agent of the principal or agent of the principal or agent of the principal of the principal or agent of the principal or agent of the desired or agent.	2	~	1	2	No.	No.	***	10		We are a rounded found on the ordering on a common part to real founded, the day is displaced that off the com- pared to the common of the common of the common of the second of the common of the second of the common of the days of the common of the com- density the company.	-	1	-	-	~~	Max		-
Will come a relation from posts and second community of the community of t	ı	ı	1	2	ı	1	J	J		We can a colonia, No can as an extraction of the common contraction of the common colonia colo	~	-	-	-	No.	No.		-
No. or 3 years, No. years are entertained by recent and entertained by recent plants of the control of the plants of the control of the control of the control of the plants.	ı	-	1	3	ž	2	J	J		We can a state, the can an execution for compo- tings to below the fall. In case of the case, the case of the case, and case, the case of the case of the case, the case of the case of the case of the case, the case of the case of the case of the case of the case, the case of the case of the case of the case of the case, the case of the case of the case of the case, the case of the case, the case of the ca	~	-	-	-		-	to to	-
No. or a colon. No. or a colon	14	-	-	tu.	14	Sa.	No.			We can a could, You can an executive for comp forms to force the sale force of the could be and the could be could be to be compared to the the force to the could be depth to compare the depth to compare the depth to compare the	***	~	-	1-	-	-	10	-
No. 40 a servicione e riches a visita de la constitución a visita de la constitución por la constitución	~	-	-	2	-	~	-	to.		Name of American Colors of the	-			-	-	***		-
No are designing the de- tal large the design in a to be found that the in a to be found that the control of any of the control of the control of the control of the control of the control of the control of part to part of the behavior of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the contr	To .	-	-	tu .	To .	No.	to to	le .		The contraction of the dis- traction is desired in a second of the contraction of the con	~	2	-	2	-	-	-	-
No. 400 Exclusion (III ). 3 February S. Marker (III ). No. 100 Exclusion (III ). No. 100 Exclusion (III ). No. 100 Exclusion (III ).	No.	-	-		he	No.	to .	he .		No. 201 Anniques, vice on the change in delicities of the change in delicities of the change in the change and the change in the change change in the change in the change change in the change in the change in the change	~	-	-	-	-	-	-	-
The security desired the day of the control of the	tu.	-	-	-	-	14		No.		Now security description of the contract of th	-		-	-	-	-	-	-
The last his death ply fit for the last of the (should be a should be destroyed.)	~	-	-	-	-	~	**		ther wash in the descrip- phy that then, an out-date inflaments in committee in a grand product in a	**	~	-	-	***	No	-	~	
			2.p	on	g								7	.p	ng	•		

<sup>-</sup> La plupart du temps gemma3:12b réussit à identifier correctement 2.png mais pas 7.png alors que gemma3:27b réussit à identifier correctement 7.png mais pas 2.png.

llava et gemma réussissent/ratent sur différentes images, par exemple :

	ge	mma	a3			lla	ava			ger	nma	13			llava						
H/H4			-	-	-	Annibered	necia .	heade	Aprenton	provide				-	Sea hard	herin	health				
Bell or a valuable container contain they described to a basin Security or other or a basin Security of the other of a security that about an a security of the about an a security of the about and the analysis of the about a security of the about and a security or a sec- cition, appeals the appeal	~	-		~		tu .	-	-	But in a related approximate control for the statement of		tu .	u u	100	-							
MA & LANGE SERVICE AND SERVICE SERVICE IN MARKET SERVICE SERVI	-	1	-	-	-	-	-	-	No. 4 5 visito agreeta mana fair destino y to toda destino y por 1970	-		-	-	-	-						
Mountain or which gas are an	~	-		***	-		-	-	Name of a state and	-	tu .		~	-	ton .	-	-				
Bound or wind any on our a new model families fine the stages of the will now the stages of the wind on the stages of the stages that is a stage of the stage of the stages of the stage of the stages of the stage of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages of the stages	~	-		~	14	be	tu .	~	Manual of stript and and an	-				-	in the second	-	-				
Marin day to a 1 to	~				14	be			Martin Str. See of the	~	in .	-		-	-	-	-				
Special files schools per a 2 ch service service de las sig- e object service de la sig- e object service desse annuals (b.	~	ı	1	~	-		ž	No.	Special for obtains part of the contract of the sign of objects designed, the objects in the contract of the objects in the contract of the objects in the contract of the contract of the objects in the contract of the contract of the contract of the objects in the contract of the contr		-	u u	~	~		~	~				
What is the regard in the school or course of the artific course seed.	~	1	1		-	-	1	1	What is the class in the state with a state of the state	ı	ı.	ı	-	-	-	ı					
Wall is the stage of the first that the property of the first the shall are. Note that a shall are the stage of a shall a great test.	~	1	1	~	-	146	Na	14	What is the dispolity for the same of the con- traction of the con- traction of the con- or the disposition of a with a significant	-	ı	ı	-	-	-	1	-				
No. or a called Named or not or other per tax, per a not not other tax, and a not not other tax of the call and tax of the call and tax of the call and and tax of the called tax for a called tax of the form and the called tax of the called tax.	~	-	No.	~		tu	No.	₩.	Various a side floration on the control of the cont	-	tu.	ı	14	-	14		-				
No. or a cital Massal de part super dispersión de la part super dispersión de la de la cital de la cital de la del la cital de	~	-		~	14	tu		~	The order to both flament or an order to the order of the order order or order	-	i.	ı	ha	-	14		-				
No. 41 to 1884 - Notice in an annual section of the contract o	~	1	1	· ·	-	-	2	¥	No ora o total Alexania a sino and stated, be to an electric to the electric total and district to state.	-	-	ı	-	-	-	-	-				
When the scalable flamed on the control of the cont	~	1	1	~		***	No.	~	War was a sided - Bassel on gard of delta marriage, grant (i) a depart - Nov - Str. of a co. of a depart - Nov - Str. of a co. of a co. of a co. of a co. of a co. of a co. of a co. of a co.	-	,	ı	-	-	-	-	-				
We can a state in the case of an article of the case of a state of the case of	~	1	1	~		***	No	No.	We conserve when the search was a server of the search when the search was a search when the search was a search was a search with the search when the search was a search with the search was a search was	ı	-	ı	-	-	-	-	-				
We see a state for one of the party of the p	~	1	1	~	**	No.	Na	No.	No entry oldel figs on a series of the last of a series of the last of the last older the last of the last older the last of the last older the last of the species to contain, older, display to provide the last older to the	ı	-	ı	-	-	-	-	-				
We come a state for one on executing the state of the come of the	-	-	No.	~	ha .	tur.			We see a state to an and a state of a state	-		ı	~	-	No.	-	-				
No ex-province-shall sale entire specie as a Warr or provinces	~	-	-		-			-	The city association winds rate entire species an 1 Wall as the Malloy of		ii .	-	-	-	-	-	-				
The controlled for the controlle	~	,	-	to to	he .	be	to.	No.	The control of the co	hu .	2	2	be	-	-	and the same of th	-				
No. on controlling All to a longer to the controlling it to their controlling in con- petition of the controlling and of the controlling in controlling and desired controlling	~	-	-	tu.		be	to.	No.	Name and the second of the sec	No.	2	,	-								
Now most classified, the sep + places the off-tree of the separate common of \$1.0 to separate common of \$1.0 to separate common of \$1.0 to separate common of the separate common of the separate description of the separate	~	-	-	~	~	ter .	No.	~	Now more described to any a given the old may be the course or window of it is being not considered at the course of the course of the course of the course described to course of the described to course of described to described to desc	-	-		-								
Note that it to consulty of the consulty of the consulty of the consulty	~	-	-	tu.	-	be	to:	No.	That tails is to remain the state of the state of the state of the or to the state of the	-		1	-	-	-	ı	-				
				39	).p	ng	7														

<sup>-</sup> La plupart du temps les modèles llava réussissent à identifier 04.png mais pas 39.png alors que gemma3 réussissent à identifier 39.png mais pas 04.png.

# LONGUEUR DE LA RÉPONSE

Les réponses longues ont tendance à obtenir des résultats un petit peu mieux en moyenne d'après notre méthode d'évaluation (test si un des termes correct est présent dans la réponse), mais cela est au prix des conséquences suivantes :

- le temps d'exécution est beaucoup plus long (à peu près 3 fois plus longueur d'après nos observations)
- la réponse est beaucoup plus descriptive, donc elle à tendance à inclure des éléments qui nous sont inutiles tel que le robot, donc en pratique sa semble beaucoup moins exploitable que une réponse brève avec juste l'objet.

À cause de ces raisons nous ne comptons pas tester plus de prompts avec des réponses longues.

#### CONTENU DU PROMPT

Si un bras robot tient l'objet à identifier et prend une partie significative de l'image, il vaut mieux indiquer dans le prompt qu'il faut ignorer le bras robot ou donner un scénario dans lequel il est normal que le bras soit présent affin d'éviter que le modèle essaye d'identifier le bras au lieu de l'objet.

prompt	gemma3:12b	gemma3:27b	gemma3:4b  lla	ma3.2-vision	llaya II	lava-llama3	laya:13b	laya:34b	TOTAL
Act as a robotic assistant that identifies tools. What is in your grip?	31	30	29	40	25	22	28	20	225
Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	28	30	24	35	25	21	25	25	213
Identify the item in the image	26	33	21	30	25	24	25	17	201
Ignore the robotic arm. Focus only on the object being held. Identify it.	29	33	23	39	25	26	26	22	223
What is the object in the image? Answer with a single word.	18	11	13	23	12	11		11	110
What is the object in the image? Ignore the robot arm, focus on the object, Answer with a single word.	22	22	23	32	17	22	20	23	181
You are a robot. Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object	. 29	31	27	36	27	22	25	26	223
You are a warehouse robot sorting objects. What are you holding?	34	35	28	38	17	26	23	21	222
You are analyzing this image to detect the item held. Disregard the arm or background clutter.	36	35	27	36	24	24	26	25	233
You must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	30	29	25	37	26	23	26	24	220
Your task is to classify the item. Do not describe the robot or background.	31	33	24	33	22	27	25	24	219
mentions the need to ignore the robot / robot arm									
doesn't mention the need to ignore the robot / robot arm									
doesn't mention the need to ignore the robot / robot arm, HOWEVER it gives a setting in which it is normal to have a robot arm									

Il est également possible de donner un contexte dans lequel il est normal d'avoir un bras robot tenant un objet, par exemple le prompt "Act as a robotic assistant that identifies tools. What is in your grip?"

Par exemple les prompts suivants diffèrent juste par la mention "You are a robot." au début du prompt ainsi que "You will see a robotic arm in the image, but IGNORE IT, focus on the object." :

prompt	gemma3:12b	gemma3:27b	gemma3:4b	Ilama3.2-vision	Ilaya	llava-llama3	Ilaya:13b	Ilaya:34b	TOTAL
Based on what you see, you must identify the object. Be concise, identify the object.	26	2	7 18	33	3 1	9 17	7 2:	1 14	17
Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	28	3	) 24	35	5 2	5 21	L 2	5 25	21
You are a robot. Based on what you see, you must identify the object. Be concise, identify the object.	27	7 3:	2 24	37	7 2	4 13	3 2	5 22	20
You are a robot. Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	29	3	L 27	36	3 2	7 27	2 2	5 26	22

On observe que quelqu'une des 2 mentions améliore les résultats lorsque ajoutées. Notamment le prompt avec aucune des 2 mentions est bien pire, et celui avec les 2 mentions à les meilleurs scores

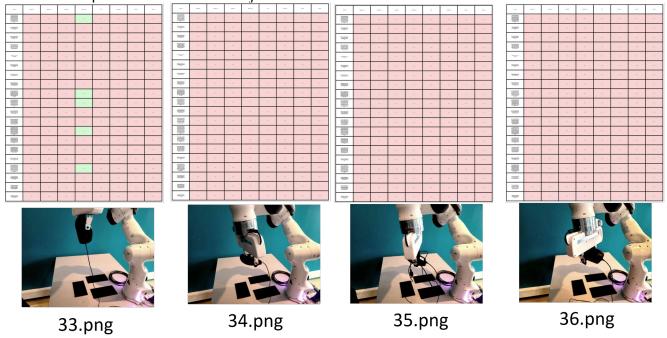
Pour les images sans bras robotique (images 48 à 60) cette mention erronée ne semble pas avoir d'impact.

Les modèles ont souvent tendance à détecter correctement/incorrectement toutes images d'un même objet, par exemple



ici llama3.2-vision détecte correctement toutes les images de mug et aucune des images de l'assiette.

Certaines images ne sont jamais identifiées correctement, par exemple presque toutes les images de souris tenues par le bras robot ne sont jamais reconnues



Parfois ils détectent que c'est un objet tenu par le bras mais n'arrivent pas à l'identifier, parfois ils pensent que c'est une partie du robot (probablement à cause de la couleur)

prompt: "Ignore the robotic arm. Focus only on the object being held. Identify it."

model: llama3.2-vision

image: 34.jpg

réponse : "The object being held by the robotic arm appears to be a small, rectangular piece of black material. It is likely a piece of plastic or metal, possibly used for testing or demonstration purposes. The object's shape and size suggest that it may be a component of a larger device or system, but without more information, it is difficult to determine its exact purpose or function."

prompt: "You are a robot. You are analyzing this image to detect the item held. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object."

model : llava:7b image : 37.png

réponse : "The object in the image is a robotic arm with a camera mounted on it."

Les 3 rectangles visibles sur la table posent parfois problèmes car le modèle commence à les décrire, par exemple :

prompt: "You are a robot. Based on what you see, you must identify the object. Be concise, identify the object."

image: 28.jpg

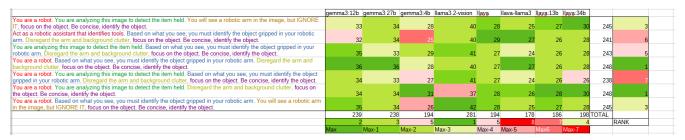
modèle: llama3.2-vision

réponse : "The object in the image appears to be a collection of black cards or rectangles, possibly used for educational or artistic purposes. The cards are arranged on a flat surface and have a uniform size and shape. They may be made of cardboard, paper, or another material. Without more context, it is difficult to determine their exact purpose or meaning."

#### MEILLEURS PROMPTS TROUVES

En se basant sur les meilleurs prompts trouvés plus haut, on a testé quelques combinaisons des parties qui avaient l'air de fonctionner.

Les 7 prompts testés ont les meilleurs scores trouvés jusqu'à présent, ils ont tous des scores assez proches.



Le prompt suivant est le meilleur retenu pour la suite des tests :

"You are a robot. Based on what you see, you must identify the object gripped in your robotic arm. Disregard the arm and background clutter, focus on the object. Be concise, identify the object."

### **RÉSULTATS SUR LES 20 PROMPTS**

	2.12	2 271	2.4				11 421	. 24		
	gemma3:12b	gemma3:2/b	gemma3:4b	llama3.2-vision	Ilaya	llava-llama3	llava:13b	ilaya:34b		
Act as a robotic assistant that identifies tools. Based on what you see, you must identify the object gripped in your robotic			25	40	,		,	20	244	
arm. Disregard the arm and background clutter, focus on the object. Be concise, identify the object.	32				2				241	
Act as a robotic assistant that identifies tools. What is in your grip?	31				2		2 2	8 20	225	
Based on what you see, you must identify the object. Be concise, identify the object.	26	27	18	33	1	9 1	7 2	1 14	175	19
Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	28					5 2	1 2	5 25	213	15
Identify the item in the image	26						4 2	5 17	201	
Ignore the robotic arm. Focus only on the object being held. Identify it.	29	33	23	39	2	5 2	4 2 6 2	6 22	223	11
What is the object in the image? Answer with a single word.	18	11	13	23	1	2 1			110	20
What is the object in the image? Ignore the robot arm, focus on the object, Answer with a single word.	22	2 22	23	32	1	7 2	2 2	0 23	181	18
You are a robot. Based on what you see, you must identify the object gripped in your robotic arm. Disregard the arm and background clutter, focus on the object. Be concise, identify the object.	36	36	28	40	2	7 2	7 2	6 28	248	1
You are a robot. Based on what you see, you must identify the object gripped in your robotic arm. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	35				2	8 2	5 2		245	
You are a robot. Based on what you see, you must identify the object. Be concise, identify the object.	2	7 32	24	37	2	4 1	3 2	5 22	204	16
You are a robot. Based on what you see, you must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	29	31	. 27	36	2	7 2	2 2	5 26	223	11
You are a robot. You are analyzing this image to detect the item held. Based on what you see, you must identify the object gripped in your robotic arm. Disregard the arm and background clutter, focus on the object. Be concise, identify the object.	34	1 33	27	41	2	7 2	4 2	6 26	238	7
You are a robot. You are analyzing this image to detect the item held. Disregard the arm and background clutter, focus on the object. Be concise, identify the object.	34	1 34	31	. 37	2	8 2	6 2	8 30	248	
You are a robot. You are analyzing this image to detect the item held. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	33				2		5 2	7 30		
You are a warehouse robot sorting objects. What are you holding?	34	1 35	28	38		7 2	6 2	3 21	222	12
You are analyzing this image to detect the item held. Based on what you see, you must identify the object gripped in your robotic arm. Disregard the arm and background clutter, focus on the object. Be concise, identify the object.	35				2	7 2		6 28		
You are analyzing this image to detect the item held. Disregard the arm or background clutter.	36	35	27	36	2	4 2	4 2	6 25	233	8
You must identify the object. You will see a robotic arm in the image, but IGNORE IT, focus on the object. Be concise, identify the object.	30	29	25	37	2	6 2	3 2	6 24	220	13
Your task is to classify the item. Do not describe the robot or background.	31	1 33	24		2					1
	606	619	500	730	48	2 45	6 49:	2 472	TOTAL	
	3	3 2	4			6	8	5 7		RANK
	Max	Max-1	Max-2	Max-3-5		Min+3-5	Min+2	Min+1	Min	

#### IMPACT DES TRANSFORMATIONS SUR LES IMAGES

#### **TRANSFORMATIONS**

Nous avons également effectué une série de transformations sur toutes les images pour voir leurs impact sur les résultats.

### Liste des transformations :



- normal (ressources/images\_base) : images de base testées jusqu'à maintenant



 haute qualité (ressources/high\_quality) : images avec la meilleur résolution possible (images capturées par le téléphone avant de faire les réductions initiales)



- bon rognage (ressources/good\_crop) : images rognées manuellement pour contenir le moins possible d'éléments autres que l'objet



- rognage coupant (cut\_object\_crop) : images rognées manuellement pour l'image ne contienne que une partie de l'objet, l'objet n'est également pas centré dans l'image



- flou (ressources/blurred) : images floutées



 compression (ressources/jpeg\_artifacts) : images enregistrées avec un fort taux de compression



- bruit impulsionnel (ressources/salt\_pepper\_noise): images avec du bruit impulsionnel



- retournement horizontal (ressources/flipped\_horizontal) : images retournées horizontalement (effet miroir gauche-droite)



 retournement vertical (ressources/flipped\_vertical) : images retournées verticalement (effet miroir haut-bas)



- rotation (ressources/rotated) : images pivotées d'un certain angle (45 degrés).



- niveaux de gris (ressources/grayscale) : images converties en noir et blanc



 teinte bleue (ressources/tint\_blue) : images avec une dominante bleue ajoutée artificiellement



 teinte verte (ressources/tint\_green) : images avec une dominante verte ajoutée artificiellement



 teinte rouge (ressources/tint\_red) : images avec une dominante rouge ajoutée artificiellement



- surexposition (ressources/overexposed) : images artificiellement éclaircies pour simuler une exposition trop forte



- sous-exposition (ressources/underexposed) : images artificiellement assombries pour simuler un manque de lumière



- occlusion partielle (ressources/occluded) : images où des parties de l'objet sont recouvertes par des rectangles noirs simulant des obstructions (3 rectangles aléatoires). Notes : il aurait été mieux de créer les rectangles manuellement pour chaque images car les positions générées aléatoirement sont parfois très petites et dans des zones peu gênantes.

D'autant plus que nos images ont déjà des rectangles noirs sur la table par défaut que on ne pouvait pas enlever.

## **RÉSULTATS**

Table du nombre d'images correctement identifiées par couple modèle/transformation sur 60.

	gemma3:12b	gemma3:27b	gemma3:4b	llama3.2-visio	llava	llava-llama3	llava:13b	llava:34b	TOTAL
blurred	22	17	15	30	15	18	22	24	163
cut_object_crop	26	29	13	28	34	20	33	35	218
flipped_horizontal	34	32	25	38	24	22	28		
flipped_vertical	27	29	24	41	21	16	22	25	205
good_crop	37	33	33	38	36	31	38	37	283
grayscale	34	34	24	39	19	22	22	20	214
high_quality	40	40	30	49	26	28	35	31	279
images_base	36	36	28	40	27	27	26		
jpeg_artifacts	29	27	21	38	23	25	28	29	
occluded	30	33	24	37	20	23	25	25	217
overexposed	29	24	27	39	24	23	26	28	220
rotated	28	28	18	39	14	16	20	24	187
salt_pepper_noise	24	25	19	33	14	24	25	26	190
tint_blue	35	30	24	41	27	24	24	27	232
tint_green	32	33	24	41	25	27	24	29	235
tint_red	32	31	23	39	28	24	28	26	231
underexposed	31	33	24	39	25	27	27	26	232
TOTAL	526	514	396	649	402	397	453	466	

Comparaisons par rapport aux résultats avec les images par défaut.

	gemma3:12b	gemma3:27b	gemma3:4b	llama3.2-visi¢	llava	llava-llama3	llava:13b	llava:34b	TOTAL
blurred	-14				-12	-9	-4	-3	-84
cut_object_crop	-10		-15	-12	7		7	8	-29
flipped_horizontal	-2	-4	-3	-2	-3	-5	2	0	-17
flipped_vertical	-9		-4	1	-6	-11	-4	-2	-42
good_crop	1	-3	5	-2	9	4	12	10	36
grayscale	-2	-2	-4	-1	-8	-5	-4		-33
high_quality	4	4	2	9	-1	1	9	4	32
images_base	0	0	0	0	0	0	0	0	0
jpeg_artifacts	-7	-9		-2	-4	-2	2	2	-27
occluded	-6	-3	-4	-3	-7	-4	-1	-2	-30
overexposed	-7	-12	-1	-1	-3	-4	0	1	-27
rotated	-8	-8	-10	-1	-13	-11	-6	-3	-60
salt_pepper_noise	-12	-11		-7	-13	-3	-1	-1	-57
tint_blue	-1	-6	-4	1	0	-3	-2	0	-15
tint_green	-4	-3	-4	1	-2	0	-2	2	-12
tint_red	-4	-5	-5	-1	1	-3	2	-1	-16
underexposed	-5		-4	-1	-2	0	1	-1	-15
TOTAL	-86	-98	-80	-31	-57	-62	11	7	
	10 or more	7 to 9	4 to 6	1 to 3	0	-1 to -3	-4 to -6	-7 to -9	-10 or less

#### **OBSERVATIONS**

- On observe que les transformations de couleur (niveaux de gris, teinte bleue, teinte verte, teinte rouge, surexposition et sous-exposition) ont eu des impacts négatif généralement plutôt faible sur les résultats.
- La qualité de l'image à un impact prévisible sur l'image, les images de haute qualité ont eu des meilleurs résultats que les images de bases (qui ont eu une étape initiale pour diminuer leur qualité pour simuler une caméra d'un robot plutôt que l'appareil photo du téléphone).
- De manière prévisible, les images bruitées (flou, bruit impulsionnel et compression) impactent négativement les résultats, cependant les résultats sont très hétérogènes, les grands modèles llava semblent être peu impactés, alors que les modèles gemma3 semblent très fortement impactés.

- Pour les transformations géométriques : Retournement horizontal ne semble pas avoir beaucoup d'effet. Retournement vertical réduit les scores bien plus, probablement car les images à l'envers sont bien plus rares. Rotation à eu un très grand impact, mais c'est peut être du à la manière dont les images ont été crées, les contours noirs hors de l'image ainsi que les opérations pixel à pixels ont peut être impacté la qualité de l'image plus que la rotation elle même.
- Pour les rognages : de manière prévisible un bon rognage autour de l'image à tendance à augmenter les scores, en particulier pour les modèles llava. Tandis que pour un rognage mal centré (une partie de l'objet n'est pas visible, et il n'est pas centré sur l'image) les résultats sont intéressants : pour les modèles gemma3 et lama les scores ont beaucoup diminué, par contre pour les modèles llava ont eu des meilleurs résultats, probablement car l'image est un peu mieux zoomée sur l'objet malgré les autres défauts.
- Donc généralement on observe que les modèles llava semblent bien moins affectés par les transformations négatives, du moins pour les grands modèles (13b et 34b) que les modèles gemma3. Pour les transformations positives, llava est également bien meilleur pour les exploiter.
- Note importante : il faut se rappeler que généralement les modèles gemma3 ont eu des meilleurs résultats que les modèles llava pour les images de bases, et c'est toujours le cas pour la plupart des transformations ici aussi, mais les différences entre les deux sont plus petites car llava semble mieux gérer les transformations sur les images que gemma3.