Analisi comparativa dei costi d’utilizzo di LLM con input immagini

# **Tabella di riepilogo**

Ipotizzando immagine dalla risoluzione massima possibile, un prompt simile a quelli dell’ultimo documento, ed una risposta ipotetica da 3000 caratteri

|  |  |  |  |
| --- | --- | --- | --- |
| **Prezzo per chiamata** | **OpenAI GPTo** | **Gemini Pro** | **Claude Sonnet 3.5** |
| **X1** | $0.023825 | €0.0025 | $0.0228 |
| **X100** | $2.3825 | €0.25 | $2.28 |

Considerazioni:

La discriminante più influente sui costi è dettata dalla lunghezza della risposta, NON dalla qualità dell’immagine.

# **Google’s Gemini Models**

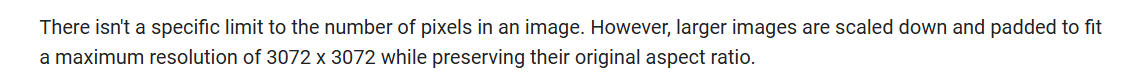
Usufruendo del modello top di gamma, Il costo ammonta a **0,0003** centesimi per una singola chiamata con immagine nell’input,

a questa cifra va sommato il costo per l’input testuale del prompt. Per dare una idea, i prompt trattati nell’ultimo documento ammontano a 2460 caratteri ergo < 1000 token.

Sommiamo ancora il costo relativo alla risposta di 0.00125 per mille catteri,

**Ipotetico costo per chiamata: Costo immagine + Costo Prompt da 3k caratteri + Costo risposta da 1k caratteri = 0.0003 + ( 3 x 0.00031 ) + 0.00125 = 0.0025 €**

Non vi è alcuna variabile dovuta alla risoluzione dell’immagine.



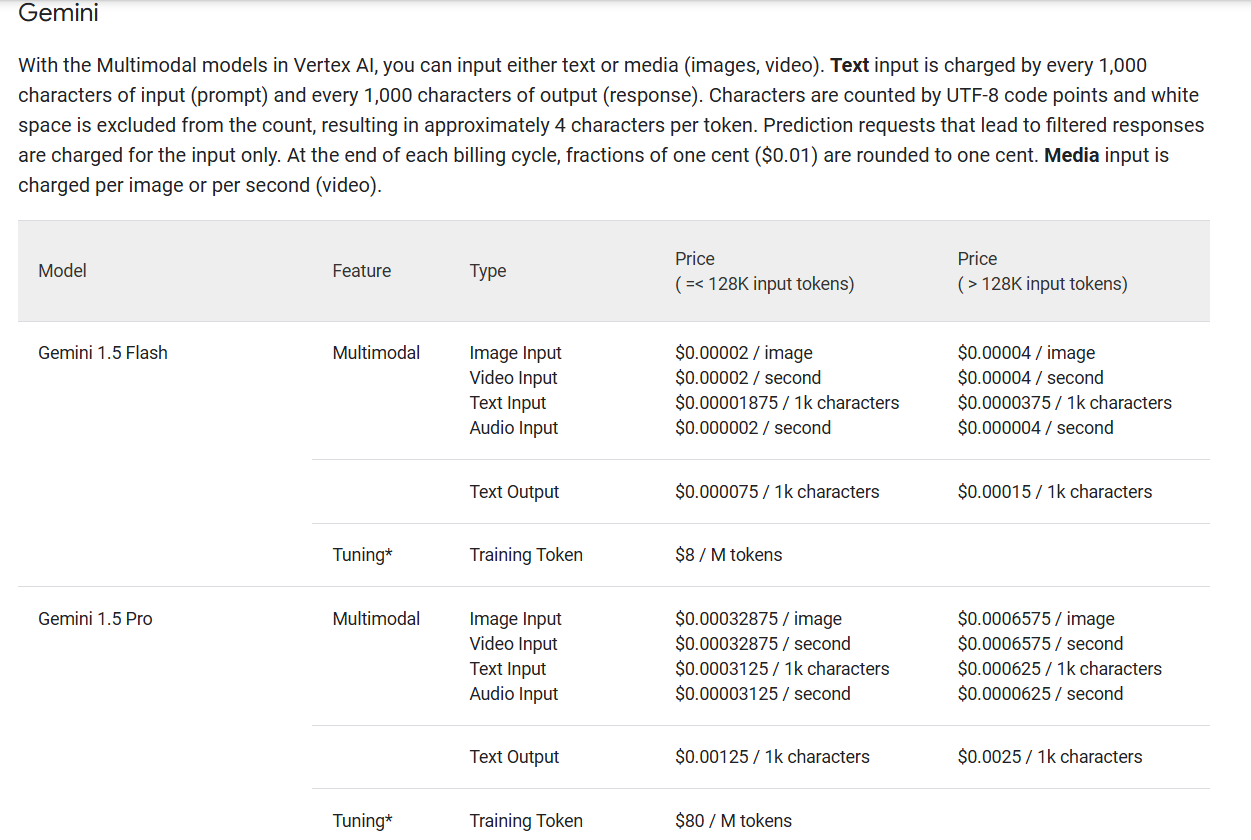
Le immagini dalla risoluzione superiore a 3072x3072 subiranno un pre processing di scaling down.

E’ ottimale utilizzare quindi immagini dalla risoluzione migliore possibile.

Fonti:

<https://cloud.google.com/vertex-ai/generative-ai/pricing#google_models>

https://cloud.google.com/vertex-ai/generative-ai/docs/multimodal/image-understanding



# **OpenAI’s GPTs**

Il costo massimo per una chiamata con immagine ammonta a: $0.003825

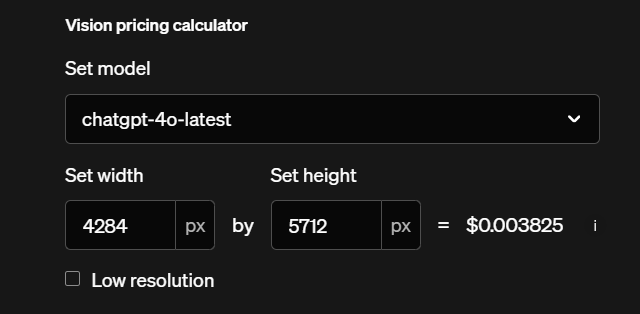
Sommando il costo di un ipotetico prompt da 1k tokens pari a $0.005 (5 dollari per milione di token)

Sommando il costo di una ipotetica risposta da 1k tokens pari a $0.015 (15 dollari per milione di token)

**Costo stimato per singola chiamata: $0.023825**

Con OpenAI le immagini saranno pre processate ad una risoluzione massima di 2048 x 2048 (mantenendo le proporzioni)

[Al seguente link](https://openai.com/api/pricing/) è possibile trovare un calcolatore di costi data la risoluzione dell’immagine in input



Fonti:

https://platform.openai.com/docs/guides/vision#calculating-costs

<https://openai.com/api/pricing/>

# **Anthropic’s Claude**

L’immagine d’input non può superare i 1600 tokens, nel senso che risoluzioni elevate saranno pre processate fino al raggiungimento di 1600 tokens.



Considerando il modello Sonnet dal costo di $3 per Milione di tokens, abbiamo un costo massimo di chiamata per immagine pari a: $0.0048

Addizionando il costo del prompt ipotetico da 1000 tokens pari a: $0.003

Addizionando ancora il costo della risposta ipotetica da 1000 tokens: $0.015 (15 euro per milione di token)

**Costo stimato per singola chiamata: $0.0228**

Fonti:

https://docs.anthropic.com/en/docs/build-with-claude/vision