

HEAD POSE ESTIMATION

Documento integrativo al fine di approfondire i risultati ottenuti

Marco Costante, Paolo La Banca, Maria Giovanna Albanese

23 giugno 2022

1 Risultati

Gli esperimenti sono stati eseguiti sul dataset BIWI. Il 70% dei fotogrammi selezionati casualmente sono stati usati per creare i modelli di riferimento e il restante 30% per la fase di test.

L'analisi è stata eseguita sull'intero dataset, i soggetti non sono stati divisi. In particolare, per ogni modello utilizzato sono stati usati gli stessi insiemi di training e test, generati attraverso lo stesso *Random state*.

I risultati ottenuti attraverso i modelli regressione sono stati analizzati con il **Mean Absolute Error (MAE)**, il quale misura la media dei valori assoluti delle differenze tra i valori predetti e quelli effettivi.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

Come possiamo vedere nell'Eq. 1 y_i è l'angolo effettivo e \hat{y}_i è l'angolo predetto.

Metodo	Pitch	Roll	Yaw	MAE
Regressione Lineare	16.40	9.43	16.49	14.10
Bayesian Ridge Regression	16.18	9.08	16.29	13.85
Lasso Regression	16.40	9.43	16.49	14.10
Gradient Boosting Regression	15.77	8.78	15.83	13.46
Extreme Gradient Regression	16.20	9.47	16.33	14.00

Tabella 1: Gradi di errori per Pitch Roll e Yaw

Metodo	Pitch	Roll	Yaw	MAE
HP ² IFS [9]	4.05	6.23	3.30	4.52
HP ² IFS-LR [10]	6.57	5.47	3.80	5.28
HP ² IFS-BRR [10]	6.57	5.46	3.80	5.28
HP ² IFS-LsR [10]	6.58	5.29	3.80	5.28

Tabella 2: Gradi di errori nello stato dell'arte per Pitch Roll e Yaw

Confrontando i valori ottenuti dal nostro approccio (Table1) con quelli presenti in letteratura (Table2), si nota come l'errore sia significativamente più alto.

Questo è dovuto al fatto che la nostra tecnica analizza tutte le possibili pose presenti nel dataset, mentre molti degli approcci pre esistenti si limitano ad un determinato range di variazione. È logico assumere che quando la posa della testa è estrema, è più complicato stimare correttamente gli angoli.

Al fine di approfondire i risultati ottenuti si mostra una rappresentazione grafica della distribuzione degli errori nella previsione degli angoli di variazioni, lungo tutti i possibili gradi 1.

Il grafico in figura 1, rappresenta nel caso della regressione lineare, come all'aumentare del valore di *pitch*, *yaw* e *roll*, aumenti anche il valore del rispettivo Mean Absolute Error.

Il grafico è stato ottenuto andando a partizionare i valori degli angoli in diverse classi, ognuna delle quali copre un range di 10 gradi (in termini di angoli di Eulero).

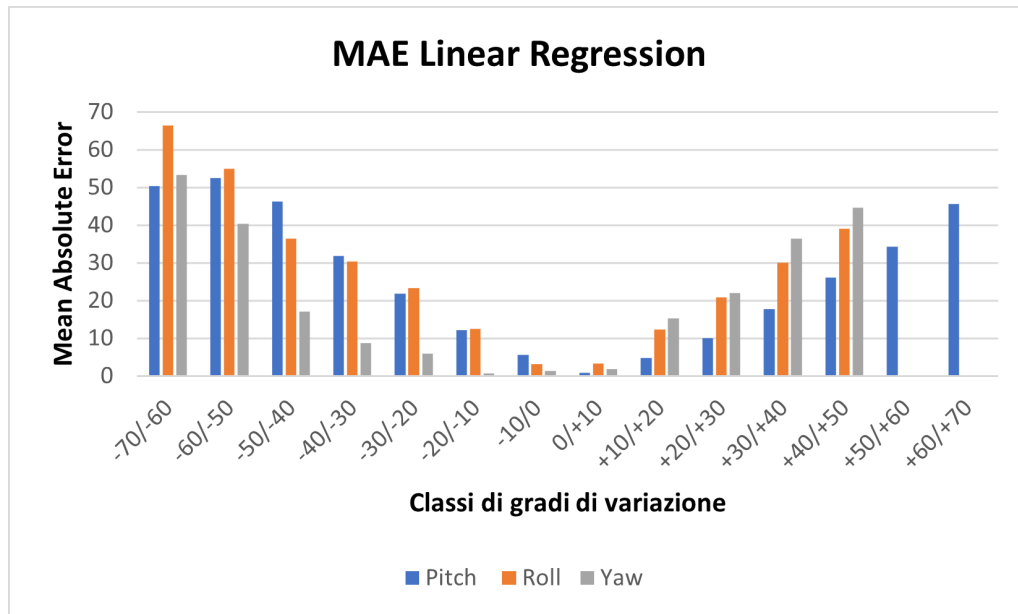


Figura 1: Distribuzione dell'errore lungo tutti i possibili gradi.

La forma del grafico a barre restituisce forte indicazione della distribuzione degli errori. Da una prima analisi risulta che la componente maggiormente caratterizzata da questo fenomeno sia *pitch*, tuttavia la maggiore variazione è data dalla componente *yaw*.

Al fine di fornire un'analisi completa vengono mostrati nelle tre sezioni successive le distribuzioni degli errori per range di variazioni, relativi al modello lineare semplice, un modello regolarizzato e un modello con gradient boosting.

1.1 Pitch

Si mostra la distribuzione della differenza tra valori reali e predetti dai vari modelli rispetto la componente *pitch*.

Gradi	Linear	Bayesian	Gradient
-70/-60	50,29	53,88	54,82
-60/-50	52,42	53,78	54,23
-50/-40	46,29	47,37	45,21
-40/-30	31,82	37,96	37,89
-30/-20	21,84	22,93	23,89
-20/-10	12,17	13,04	13,67
-10/0	5,57	6,19	6,17
0/+10	0,76	1,05	1,49
+10/+20	4,80	5,03	4,70
+20/+30	10,95	10,63	10,80
+30/+40	17,77	18,72	19,67
+40/+50	26,13	27,49	27,45
+50/+60	34,21	35,78	37,49
+60/+70	45,63	46,69	47,78

Tabella 3: MAE per pitch e gradi di variazione a confronto

1.2 Yaw

Si mostra la distribuzione della differenza tra valori reali e predetti dai vari modelli rispetto la componente *yaw*.

Gradi	Linear	Bayesian	Gradient
-70/-60	53,34	54,95	59,82
-60/-50	40,37	42,17	49,34
-50/-40	17,13	27,31	29,69
-40/-30	8,73	17,98	18,67
-30/-20	5,99	9,50	9,49
-20/-10	0,69	6,61	6,48
-10/0	1,34	2,12	3,13
0/+10	1,75	0,77	1,74
+10/+20	15,20	1,11	0,26
+20/+30	22,07	2,32	4,62
+30/+40	36,38	10,03	11,04
+40/+50	44,55	15,98	24,42
+50/+60	0	22,88	39,52
+60/+70	0	37,33	0

Tabella 4: MAE per yaw e gradi di variazione a confronto

1.3 Roll

Si mostra la distribuzione della differenza tra valori reali e predetti dai vari modelli rispetto la componente *roll*.

Gradi	Linear	Bayesian	Gradient
-70/-60	66,45	64,76	61,14
-60/-50	54,88	54,15	53,90
-50/-40	36,37	38,92	40,23
-40/-30	30,42	31,37	31,83
-30/-20	23,26	23,91	24,76
-20/-10	12,53	12,99	8,43
-10/0	3,12	3,22	3,08
0/+10	3,28	3,44	3,69
+10/+20	12,37	12,96	13,12
+20/+30	20,83	21,71	21,70
+30/+40	30,04	31,73	23,98
+40/+50	38,98	0	40,72
+50/+60	0	0	0
+60/+70	0	0	0

Tabella 5: MAE per roll e gradi di variazione a confronto

2 Approfondimenti

Al fine di avvalorare la tesi illustrata nel paragrafo precedente, è stato effettuato un ulteriore studio, andando a valutare le prestazioni dei modelli di regressioni precedentemente descritti, su una porzione di fotogrammi caratterizzati da un range di variazione limitato.

In particolare, sono stati selezionati esclusivamente i soggetti con range di variazione compreso tra -40° e 40° .

Come previsto i risultati restituiti dai modelli di regressione sono stati nettamente migliori 6.

Metodo	Pitch	Yaw	Roll	MAE
Regressione Lineare	12.15	10.6	6.36	9.70
Bayesian Ridge Regression	11.9	10.44	6.01	9.45
Lasso Regression	12.15	10.64	6.36	9.71
Gradient Boosting Regression	11.91	10.67	6.23	9.60
Extreme Gradient Regression	11.49	10.27	5.78	9.18

Tabella 6: Gradi di errori per Pitch Roll e Yaw su range ristretto

Chiaramente, anche in questo caso, gli errori più alti saranno concentrati maggiormente in prossimità degli angoli estremi, come osservabile dalla figura 2.

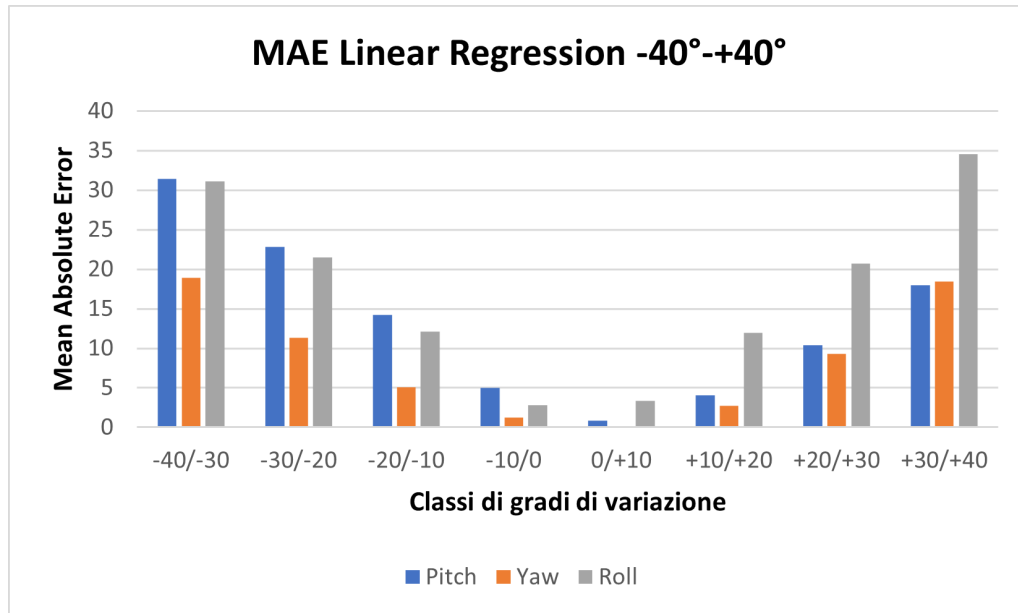


Figura 2: Distribuzione dell'errore lungo tutti i possibili gradi.

Si nota in questo caso che le componenti soggette maggiormente influenzate dall'errore siano *pitch* e *roll*.

Il trend generale della distribuzione degli errori in termini di percentuale delle immagini testate può essere visto in figura 3.

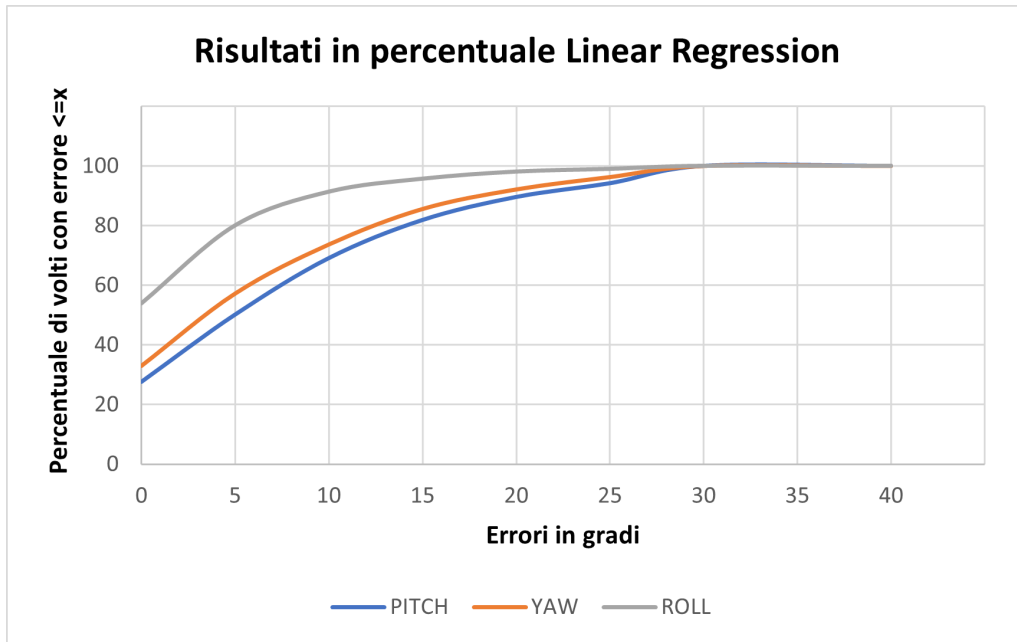


Figura 3: Errori in termini di percentuale delle immagini testate.

I risultati mostrano che:

- *Pitch* risulta non avere errori per il 27,54%, il 50,11% dei dati ha un errore tra 0 e 5 e il 69,06% per errori inferiori a 10 gradi.
- *Yaw* risulta non avere errori per il 32,97%, il 57,21% dei dati ha un errore tra 0 e 5 e il 73,66% per errori inferiori a 10 gradi.
- *Roll* risulta non avere errori per il 53,84%, il 80,01% dei dati ha un errore tra 0 e 5 e il 91,33% per errori inferiori a 10 gradi, confermandosi la componente con i risultati migliori.

3 Conclusioni

Alla luce di questa analisi dei risultati è stato mostrato come un elevata angolazione della testa influenzi fortemente l'accuratezza dei risultati di previsioni. Si mostra come selezionando un range ristretto si ottengano risultati migliori, ma comunque maggiori di altri approcci presenti in letteratura.

Si potrebbe pensare che ciò sia dovuto ai valori definiti nell'esecuzione dell'algoritmo di codifica frattale, assumendo come per valori maggiori di range e d'angolo si potrebbero ottenere previsioni più accurate.