Rubric-Constrained Figure Skating Scoring

Arushi Rai and Adriana Kovashka

University of Pittsburgh | School of Computing and Information

Introduction

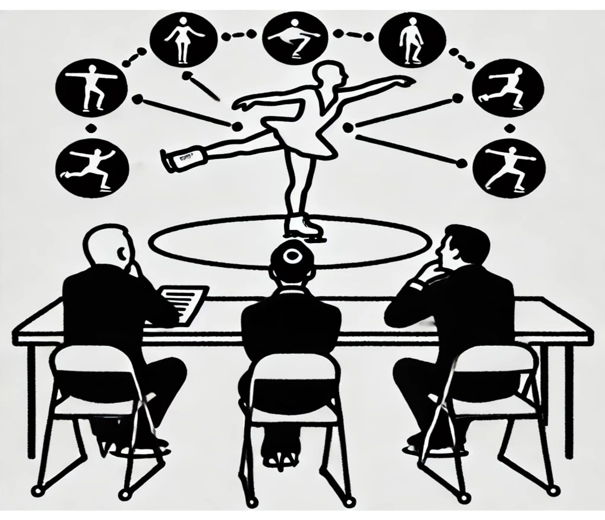


Image generated from DALL-E 3.

Figure. Figure skating judges use detailed rubrics to score each element in a figure skating routine.

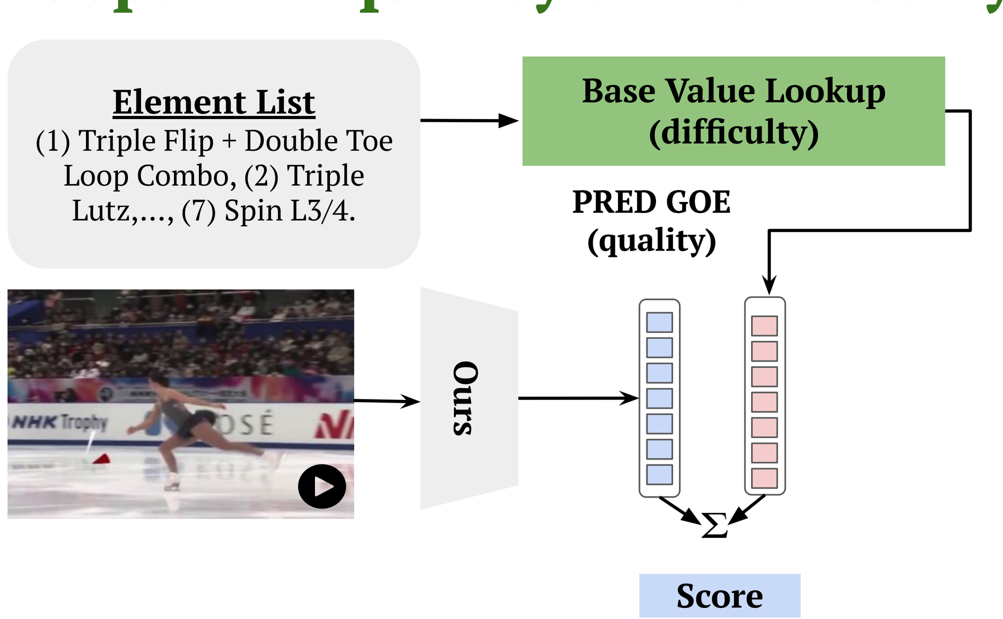
How can we utilize existing rubrics that are applied to each element?

Two challenges from existing datasets (Fis-V and FS1000):

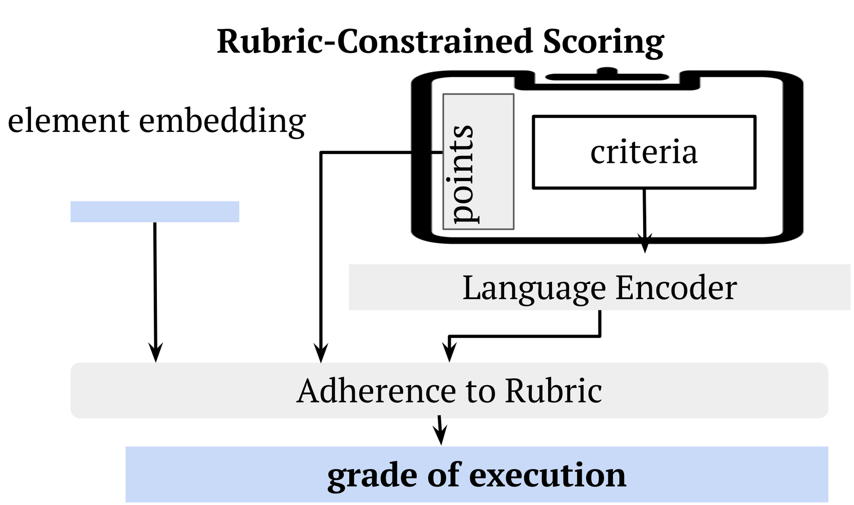
1. contain video-level scores which entangle both difficulty and quality, but (a) the rubric only assesses quality and (b) quality is more important than difficulty
2. no element-segmentations

We propose an interpretable method that uses rubric information and computes element-level scores without additional annotation effort.

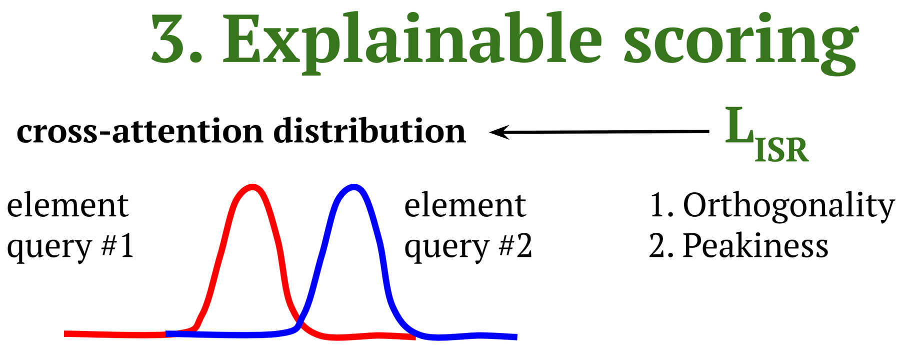
1. Separate quality and difficulty



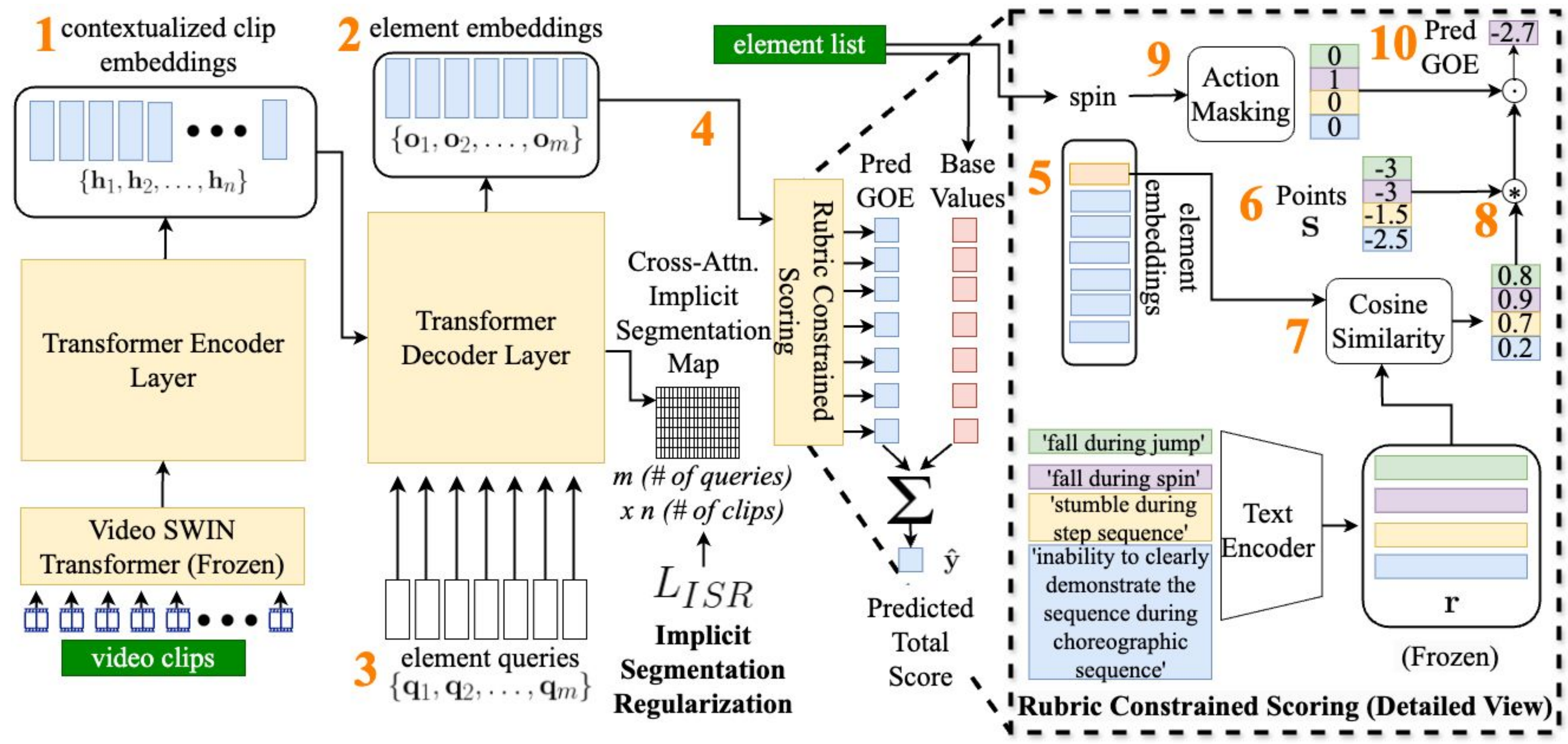
2. Interpretable scoring



3. Explainable scoring



Method



Implicit Segmentation Regularization

A is the cross-attention distribution between queries and contextualized clip embeddings.

Orthogonality loss $L_o = \sum_{k=0}^{m}\sum_{j=0}^{m} \mathbf{A}_k \cdot \mathbf{A}_j \quad \text{for} \quad k \neq j$

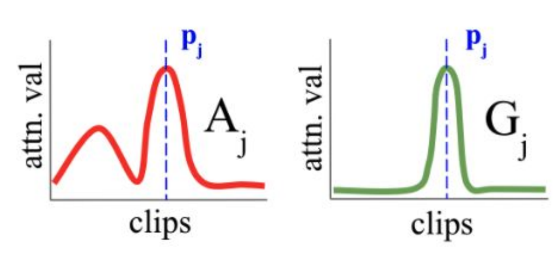


Figure. Desired attention distribution ($\mathbf{G}_j$) is greedily generated from the actual attention distribution ($\mathbf{A}_j$) for a given query.

Peak loss $L_p = \sum_{j=1}^{m} D_{\mathrm{KL}}(\mathbf{A}_j \| \mathbf{G}_j)$

Rubric-Constrained Scoring

r are the text embeddings computed for each rubric item. **o** are element embeddings for a video *i*.

$$w_{ij} = s_k \cdot \cos(\mathbf{o}_i^j, \mathbf{r}_k)$$

$$\mathrm{GOE}_i^j = 6 * \sigma(\sum_k s_k \cdot \cos(\mathbf{o}_i^j, \mathbf{r}_k)) - 3$$

BVL is the Base Value Lookup which stores a mapping between element names to difficulty values.

$$\hat{y}_i = \sum_{j}^{m} BVL(j) + GOE_i^j$$

$$L_{se} = (y_i - \hat{y}_i)^2$$

Pretraining

1. Visual-Text Pretraining
2. Visual-Only Pretraining
3. Joint Pretraining

Experiments

Method	MSE (↓)	Sp. Corr. (↑)
CoRe** [25] (2021)	23.50	0.66
GDLT* [23] (2022)	33.60	0.69
TPT** [1] (2022)	27.50	0.57
MLP-Mixer** [21] (2023)	19.57	0.68
SGN [6] (2024)	<u>19.05</u>	0.70
Base Value Lookup (BVL)	19.53	0.76
GDLT (2022) [23] w/ BVL	28.52	<u>0.77</u>
Ours	**9.34**	**0.84**
GT Base Value	12.03	0.91

Table 1. Scoring evaluation on Fis-V.

1. Prior methods don't outperform using difficulty scores
2. Our method significantly improves over SoTA in both score precision and ranking.

Method	Order-Insensitive Precision (1:1 Assignment) (%)
Ours w/o $\mathcal{L}_{ISR}$	12.5
+ $\mathcal{L}_{peak}$	3.6
+ $\mathcal{L}_{ortho}$	32.1
+ $\mathcal{L}_{ISR}$	33.9
+ $\mathcal{L}_{ISR}$, PT_{vis}	37.5
+ $\mathcal{L}_{ISR}$, PT_{vt}	42.9
+ $\mathcal{L}_{ISR}$, PT_{joint}	35.7

Table 2. Implicit segmentation evaluation on Fis-V test subset.

Implicit segmentation regularization improves the ability of queries to attend to actual elements over background transitions

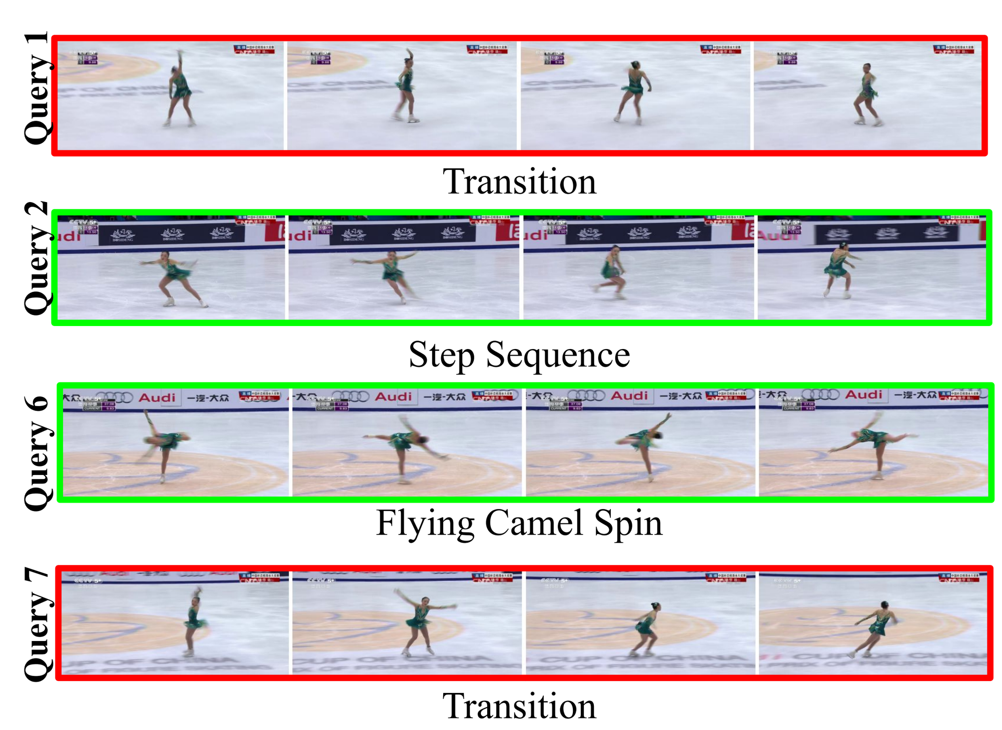


Figure. Score prediction on a video from Fis-V. Half of queries attend to transitions which are difficult to discern from elements.

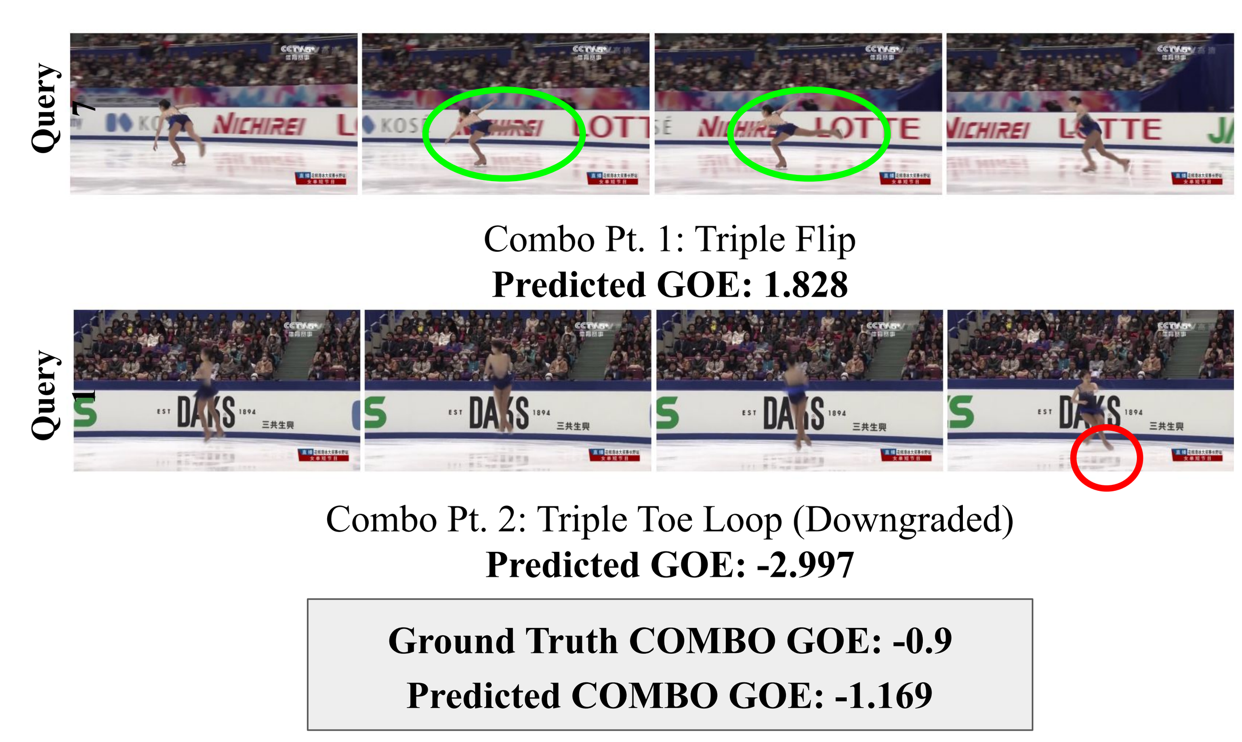


Figure. Edge case - two queries might attend to the same element, however, this might be due combination jumps.

Conclusion

We showed a new, interpretable, well-performant mechanism for action quality assessment in figure skating. It relies on a well-defined rubric of criteria for figure skating elements, and an implicit segmentation approach to obtain element-level scoring. Our approach is a first step in using freely available, structured, language-based resources for improving interpretability in figure-skating scoring.