Hierarchical NeuroSymbolic Approach for Comprehensive and Explainable Action Quality Assessment—Supplementary Material

Lauren Okamoto
Princeton University
lokamoto@princeton.edu

Paritosh Parmar
IHPC, A*STAR, Singapore
paritosh.parmar@alumni.unlv.edu

Annotated training data and code: https://github.com/laurenok24/NSAQA.

1. Neural Network Training Implementation Details

1.1. Platform Detection

We used the Detectron2 model developed by Facebook AI Research to perform object detection to classify the platform in a given frame [6]. We trained detectron2 on a new dataset composed of 854 random images from the FineDiving dataset [7], spanning a wide variety of diving venues, camera distances/angles, types of dives, and parts of the dive. Each image is annotated with the platform (if visible), and random other parts of the frame (i.e. the diver, background, audience, etc.) labeled as "non-platform." We trained detectron2 to detect two classes of platform and non-platform over 700 epochs.

1.2. Diver Detection and Pose Estimation

Detecting the diver and what pose they're in is split into two steps: (1) detecting the diver using Detectron2 [6], then (2) performing pose estimation. The first step of detecting the diver was done similarly to the platform and splash detection. We constructed a new dataset of 920 diverse images from the FineDiving dataset annotated with the diver in the frame.

1.3. Diver Pose Estimation

In order to do pose estimation on the diver, we fine-tuned a human pose estimation model by Sun et al. called High Resolution Net (HRNet) that maintains high resolution representations throughout the whole process [5]. We used the ExPose dataset to fine-tune HRNet [4]. The ExPose dataset contains 3000 diving images together with their annotations in the MPII dataset format [1]. The images were from four different individual diving events on both springboard and platform. We combined this dataset with 2570 diving images from the 2012 Olympics, totaling to 5570 annotated images. We use the same training parameters as what was used to train the base HRNet model: Adam optimizer [3] and a base learning rate of 1e-3 (1e-4 and 1e-5 at 170th and 200th epochs respectively). We train over 210 epochs.

1.4. Splash Detection

Splash detection was trained in essentially the same way as the platform. We trained the detectron2 model [6] on a new dataset with 998 random images sampled from the FineDiving dataset [7]. These images, like the platform samples, are diverse across venues, camera points of view, dives, divers, and parts of the dive. Each image is annotated with the splash (if visible), and random other parts of the frame (i.e. the diver, background, audience, platform, etc.) labeled as "non-splash." We trained detectron2 to detect two classes of splash and non-splash over 700 epochs.

2. Action Recognition Microprogram Algorithms

2.1. Somersault Counter

We count the number of somersaults performed in a given dive by looking at the vector from the pelvis position to the thorax position. We track the rotations of that vector to count the somersaults performed by a diver. A dive generally starts with the pelvis-to-thorax vector pointing upwards at 90° (if the diver is standing) or downwards at negative 90° (if the diver is in an armstand). In diving, a half somersault of ideally 180° is considered to be performed to a minimal extent so long as the

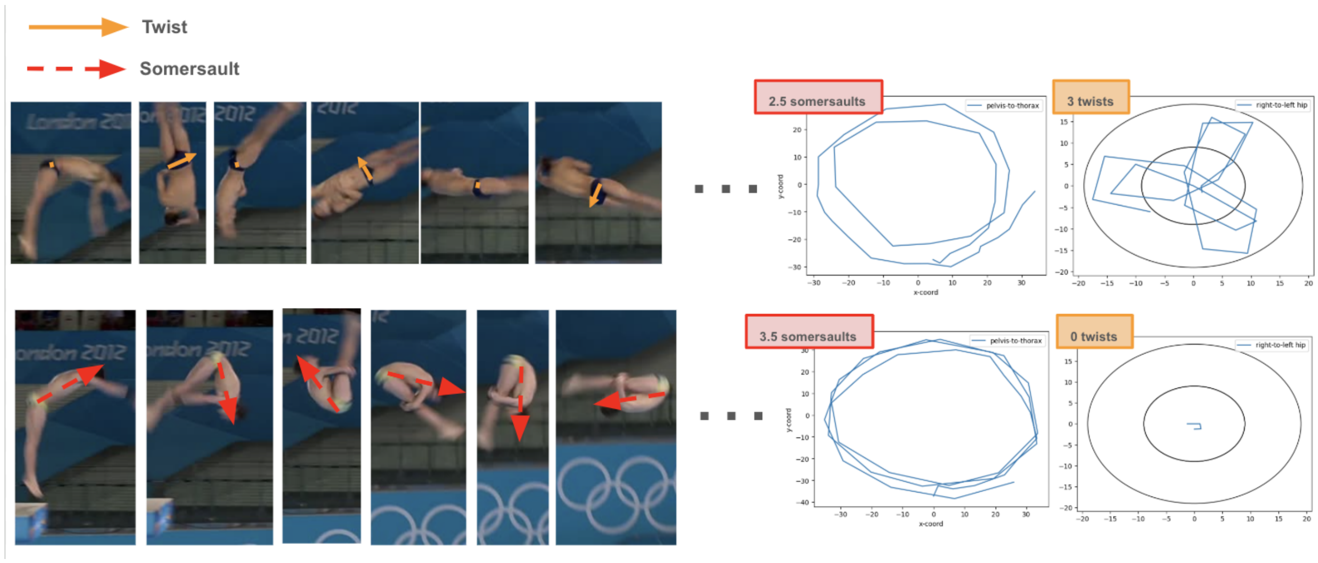


Figure 1. Somersault and Twist Counters applied to two example dives. **First row**: a forward dive with 2.5 somersaults (2.5 revolutions in the left diagram) and 3 twists (6 "petals" in the right diagram). **Second row**: a forward dive with 3.5 somersaults (3.5 revolutions in the left diagram) in pike position and no twists (zero "petals" in the right diagram).

diver rotates a majority of the 180°. For example, if a diver rotates 130° of the 180°, they will be given credit for the half somersault such that she will not fail the dive. We credited half somersaults as long as the diver rotated 110° of the 180°.

2.2. Twist Counter

We count the number of twists performed in a given dive by looking at the vector from right hip to the left hip. As a diver makes a 180° turn or twist of their body, the trace of this vector forms a "petal" shape. This is because when the diver is in profile in the image frame, the left and right hips are effectively overlapping in position such that the vector is near the origin. As the twisting rotates the diver to a frontal (or rear) view, the left and right hips become separated in the image, until a maximum separation is reached at near direct frontal (or rear) view, which corresponds to 90° of twist. Subsequent twisting rotates the diver back to a profile view where separation is again minimal. This position indicates completion of a 180° twist from the previous time that the diver was in profile view. In our plot of the right-to-left hip vector, such a 180° twist is seen as one petal shape. In our implementation, we count a petal (180° twist) by tracking when the right-to-left hip vector leaves and returns to a region centered around the origin. Generally, a diver starts in profile such that the vector is near the origin. When the vector leaves the central region and extends to the outer region, this indicates that the diver has rotated 90° and is now in a front or rear view in the image. When the vector returns to the central region, this indicates that the diver is again in profile view and hence a 180° or half twist has been completed. Thus, the number of round trips from center region to outer region and back is equal to the number of half twists performed.

2.3. Rotation Type

There are four different rotation types: front, back, reverse, and inward. We determine rotation type by utilizing two pieces of information: (1) the direction in which the diver rotates (clockwise or counterclockwise) and (2) the direction they're facing (forwards or backwards).

To determine whether the diver is rotating clockwise or counterclockwise, we track the pelvis-to-thorax vector across frames (same as the somersault counter).

To determine whether the diver is facing forwards or backwards, we look at the knee bend of the diver. Assuming that the human knee can only bend to a large degree in one direction, we calculate the direction of rotation from the knee-to-hip vector to the knee-to-ankle vector.

2.4. Position

To determine which position the diver is in (straight, pike, tuck, or free), we looked at the degree of knee bend throughout the dive. If the knee bend (angle between the knee-to-hip vector and knee-to-ankle vector) was less than 60º for a significant portion of the dive, the diver was in the tuck position. If the diver performs a dive with twists and does less than 2.5 somersaults, the diver was in the free position. Otherwise, the diver was in the pike position.

2.5. Armstand?

We determine if a diver is doing an armstand dive by looking at the beginning of the dive, and seeing whether the thorax is below the pelvis or not. If the thorax is below the pelvis, then the diver is in the armstand position. Otherwise, they are not.

3. Temporal Segmentation Microprogram Algorithms

3.1. Start/takeoff Phase

Identifying the start/takeoff phase relies on the pose estimation of the diver and the detection of the platform. Simply, if the diver is on the platform and has not jumped below the platform, we say that the diver is in the start/takeoff phase. In order to determine whether or not the diver is on the board, we compare the diver's position in relation to the end of the platform. If the diver's hips are past the end of the platform and their ankles are high enough off the platform to count as a jump (wrists if a handstand dive), then the diver is considered as no longer on the platform and so past the start phase. As implemented, a diver is considered as still in the start phase if the x-coordinate of the hip is not past the x-coordinate of the edge of the platform or if the distance between the ankle and the end of the platform is less than 1.5 times the distance between the ankle and the knee, we can be confident that the diver is still in the start/takeoff phase.

3.2. Somersault Phase

Determining whether the diver is in the somersault phase has the following criteria. If the diver is still on the board or has completed the expected number of somersaults already, we determine that the diver is not in the somersault phase. If the diver hasn't done the expected number of somersaults, then we look at hip bend to determine whether the diver is in the somersault phase or in the twist phase. We say the frame is in the somersault phase only if the hip bend is less than 80 degrees.

3.3. Twist Phase

The twisting phase is similar to the somersault phase. If the diver is on the board or the diver has completed the expected number of twists, then the diver is not in the twisting phase. If the hip bend is greater than 80 degrees and the diver hasn't completed the expected number of twists, then the diver is in the twist phase.

3.4. Entry Phase

The diver is in the entry phase if a splash is detected and/or if the diver is no longer detected. Additionally, the diver is also in the entry phase if the expected number of somersaults and twists have been completed.

4. Detailed Dive Quality Assessment Algorithms

4.1. All Dive Elements with Descriptions + Visualizations

See Table 1

4.2. Distance From Platform Algorithm

We calculate the minimum distance of any body part from the edge of the platform. This distance is normalized by the span between the diver's thorax and pelvis joints to account for different camera distances.

4.3. Height Off Platform Algorithm

We calculate the maximum vertical height of any body part from the edge of the platform. Like the Distance From Board algorithm, the height is normalized by the span between the diver's thorax and pelvis joints to account for different camera distances.

4.4. Feet Apart Algorithm

We calculate the angle between the vectors pointing from a midpoint between the diver's knees to the left and right ankles. This angle is averaged over the duration of the dive after takeoff. The average midpoint-ankles angle is ranked against corresponding average angles from all platform dives in the FineDiving dataset (with smaller average angle being better).

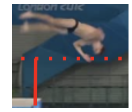
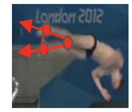
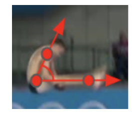
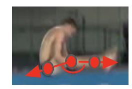
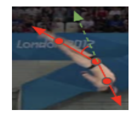
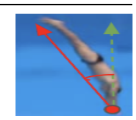
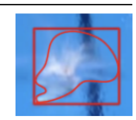
Distance-from-platform. The closest distance a diver comes to the platform during their dive is important for two reasons. First, there is a safety concern as a diver can be severely injured or even die by hitting the concrete platform. Hence, a diver is penalized for approaching too close to the platform on the way down. Second, being too far from the platform results in an unaesthetic trajectory and is also penalized.	
Height-off-platform. The higher a diver is able to jump off the platform, the more time they will have to complete their dive. Additionally, the higher trajectory of the diver off the platforms allows for an aesthetic takeoff. A diver is generally penalized for having a low height off the platform.	
Feet-apart. A streamlined and graceful form throughout a dive is considered ideal. Taking a cue from gymnastics, one aspect of an ideal diving form is maintaining one's feet together. A diver is penalized for having their feet apart.	
Somersault tightness. During a somersault, it is desirable for a diver to have as tight a tuck or pike position as possible. A very tight somersault position allows a diver to spin faster and is considered ideal.	
Knee straightness. Unless in the tuck position, the diver should keep their legs as straight as possible during flight. Bend legs results in an unaesthetic position.	
Twist straightness. During a twist, a diver should have a straight body position from head to toe. A bent body results in an unaesthetic and slower twist.	
Over/under-rotation. Each particular dive has a specified number of rotations of somersaults to be performed. The number of rotations is in half-rotation increments. A diver preferably performs the exact number of rotations (to the degree) that are specified for the dive–no more and no less. Over rotation and under rotation are penalized.	
Straightness of body during entry. During entry into the pool, a diver should ideally have a straight body form. This is aesthetically pleasing and shows that the diver maintains strength in their core. It also enables a less splashy entry.	
Splash size. Ideally, a diver has little or no splash upon entry into the water. Little or no splash after a dive from 10 meters high is difficult to achieve and is indicative of very good form upon entry.	

Table 1. **Elements of dive quality, their descriptions and visualizations.**

4.5. Somersault Tightness Algorithm

To calculate the hip bend, we measure the angle between two vectors– one pointing from the diver's hip to their thorax, and the other pointing from the diver's hip to their knee. The closer the angle is to 0°, the better. This angle is averaged over the duration of the somersault.

4.6. Knee Straightness Algorithm

To calculate the knee bend, we measure the angle between two vectors– one pointing from the diver's knee to their hip, and the other pointing from the diver's knee to ankle. The closer to 0°, the better. This angle is averaged over the duration of the somersault.

4.7. Twist Straightness Algorithm

We calculate how much the hip bend angle deviates from the ideal 180°. The angle of deviation is averaged over the duration of the twisting.

```python
"""
    Entry Phase Microprogram

    Parameters:
    - filepath (str): file path where the frame is located
    - above_board (bool): True if the diver is above the board at this frame
    - on_board (bool): True if the diver is on the board at this frame
    - pose_pred: pose estimation of the diver at this frame (None if no diver detected)
    - expected_twists (int): number of twists in full dive (from action recognition)
    - petal_count (int): number of twists counted by this frame
    - expected_som (int): number of somersaults in full dive (from action recognition)
    - half_som_count (int): number of somersaults counted by this frame
    - frame: Pil Image of frame
    - splash_detector: splash detector model
    - visualize: True if you want to save the splash segmentation mask prediction to an image
    - dive_folder_num: if visualize is true, this is where the image will be saved

    Returns:
    - 0 if frame is not in entry phase
    - 1 if frame is in entry phase
"""
def entry_microprogram_one_frame(filepath, above_board, on_board, pose_pred, expected_twists, petal_count,
expected_som, half_som_count, frame=None, splash_detector=None,  visualize=False, dive_folder_num=None):
    if above_board:
        return 0
    if on_board:
        return 0
    splash = get_splash_from_one_frame(filepath, im=frame, predictor=splash_detector, visualize=visualize,
dive_folder_num=dive_folder_num)
    if splash:
        return 1
    # if completed with somersaults, we know we're in entry phase
    if not expected_som > half_som_count:
        return 1
    if expected_twists > petal_count or expected_som > half_som_count:
        return 0
    return 1
```

```python
"""
    Feet Apart Microprogram

    Parameters:
    - filepath (str): file path where the frame is located
    - pose_pred: pose estimation of the diver at this frame (None if no diver detected)
    - diver_detector: diver detector model
    - pose_model: pose estimation model (HRNet)

    Returns:
    - feet-apart angle for given frame if exists
    - None if no diver is detected
"""
def applyFeetApartError(filepath, pose_pred=None, diver_detector=None, pose_model=None):
    if pose_pred is None and filepath != "":
        diver_box, pose_pred = get_pose_estimation(filepath, diver_detector=diver_detector,
pose_model=pose_model)
    if pose_pred is not None:
        pose_pred = np.array(pose_pred)[0]
        average_knee = [np.mean((pose_pred[4][0], pose_pred[1][0])), np.mean((pose_pred[4][1], pose_pred[1]
[1]))]
        vector1 = [pose_pred[5][0] - average_knee[0], pose_pred[5][1] - average_knee[1]]
        vector2 = [pose_pred[0][0] - average_knee[0], pose_pred[0][1] - average_knee[1]]
        unit_vector_1 = vector1 / np.linalg.norm(vector1)
        unit_vector_2 = vector2 / np.linalg.norm(vector2)
        dot_product = np.dot(unit_vector_1, unit_vector_2)
        angle = math.degrees(np.arccos(dot_product))
        return angle
    else:
        return None
```

Table 2. Examples of Microprograms for Entry Phase and Feet Apart implemented in Python. Note that these examples are just for reference, we will opensource all of our code at a later date.

4.8. Over/Under-Rotation Algorithm

To measure this, we calculate the angle from the vertical vector [0, 1] and the vector pointing from the diver's head to the ankle. A 0° angle indicates a vertical entry.

4.9. Straightness of Body During Entry Algorithm

We calculate the hip bend angle at entry, and see how much it deviates from the ideal 180°. The smaller the deviation angle, the better.

4.10. Splash Size Algorithm

We calculate the area of the detected splash and scale it to the distance between the thorax and pelvis squared to roughly account for variation in the camera distance.

5. Examples of Microprograms implemented in Python

See Table 2.

6. Visio-Linguistic Report Generation

We automatically generate a summary report of the scores calculated for each error of the dive accompanied with visual feedback of where the error occurred in the dive. Once we calculate the overall score and percentile scores of each error, we use an HTML template with pre-written language to generate the report for the dive. The language used throughout the report was crafted by a domain expert to be easily understood by the diving community. See an example report in Figure 2

6.1. Overall Score

The report starts by saying: "Your dive was __*insert score category*__, and scored a __*insert overall score*__. Here is how we scored each component of the dive. Each percentile is relative to dives in the semifinals and finals of Olympic and World Championship competitions."

The score category refers to where the overall score lies in Table 3 by World Aquatics (AQUA).

6.2. Error Table

The report then goes into each error in detail in the form of a table. The table provides the following information for each error:

- **Error:** the name of the dive error.
- **Description:** a description of how well the diver performed with respect to that particular error. The description will contain specific details for what was detected and where that lies relative to other Olympic and world championship dives in the form of a percentile.

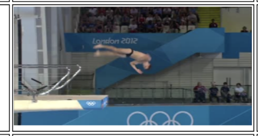
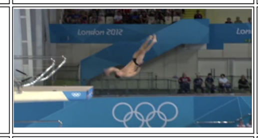
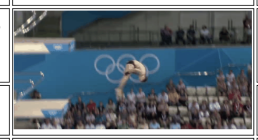
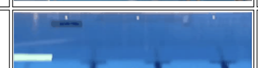
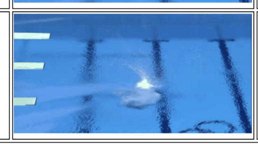
Your dive was satisfactory, and scored a **5.7**. Here is how we scored each component of the dive. Each percentile is relative to dives in the semifinals and finals of Olympic and World Championship competitions.			
Error	**Description**	**Visuals**	**Score**
Feet Apart	We found that your leg separation angle was on average 5 degrees for your dive. This is rated as 22 percentile.		2.2
Height off platform	Your jump was a bit on the lower side, and was rated as 27 percentile. Here is the highest you jumped off the platform.		2.7
Distance from platform	You were safe, but too far from the platform. Here is where you came closest to the platform.		Too Far
Somersault tightness	We found that the tightness of your pike was 54 degrees on average. This is rated as 41 percentile. Here are some examples of your position in the somersault.		4.1
Knee straightness	We found that your knees bent 9 degrees on average. This is rated as 78 percentile.		7.8
Twist Straightness	We found that the tightness of your pike was 54 degrees on average. This is rated as 41 percentile. Here are some examples of your position in the somersault.		5.5
Verticalness (over/under rotation)	We found that you deviated from vertical by 6 degrees, which was the 63 percentile.		6.3
Body straightness during entry	The straightness of your body during entry deviated by 11 degrees, which was the 82 percentile.		8.2
Splash	Your splash was small and rated as 95 percentile.		9.5

Figure 2. Example of a Visio-Linguistic Report. Visuals are shown as still images here, but are GIFs when generated as an HTML file.

Excellent	10
Very Good	8.5-9.5
Good	7-8
Satisfactory	5-6.5
Deficient	2.5-4.5
Unsatisfactory	0.5-2
Completely Failed	0

Table 3. World Aquatics Scoring Table [2]. This table shows the scores and their categories when scores are made in 0.5 point increments. If the overall score was a 9.0, the score category would be "very good". If the overall score was a 2.5, the score category would be deficient.

- **Visuals:** the visuals of each error will be in the form of a GIF highlighting the part of the dive that the system was looking at when scoring that error.
- **Score:** the percentile score of that error (divided by 10 to be more consistent with the diving scoring system which is out of 10).

6.2.1 Feet Apart Error

- **Error:** "Feet Apart"

- **Description:** "We found that your leg separation angle was on average __*insert average feet apart angle*__ degrees for your dive. This is rated as __*insert feet apart percentile*__ percentile." The average feet apart angle and its corresponding percentile score are passed in as parameters when generating the report using the HTML template.
- **Visuals:** GIF highlighting the frames where the feet apart angle peaked. The GIF is generated before rendering the HTML template, and passed in as a parameter to be displayed in the report. If there were no peaks, then the text shown is "There were no particular instances to show where your feet came apart."
- **Score:** feet apart percentile score scaled to be out of 10 points.

6.2.2 Height off Platform Error (not applicable if the dive is an armstand dive, in which case will not be included in the report)

- **Error:** "Height Off Platform"
- **Description:** "Your jump was __*insert either "good" if percentile was above 50% or "a bit low" if lower than 50%tile*__, and was rated as __*insert height off platform percentile score*__ percentile. Here is the highest you jumped off the platform."
- **Visuals:** Displays the frame where the diver had the highest height detected.
- **Score:** height off platform percentile score scaled to be out of 10 points.

6.2.3 Distance from Platform Error

- **Error:** "Distance from platform"
- **Description:** "You were __*insert 'a good distance from', 'too far from', or 'too close to'*__ the platform. Here is where you came closest to the platform.
- **Visuals:** Displays the frame where the diver was the closest to the platform.
- **Score:** "Good" if good distance, "Too far" if diver was too far, or "Too close" if diver was too close to the platform

6.2.4 Somersault Tightness Error (not applicable if no somersault phase of the dive is detected, in which case will not be included in the report)

- **Error:** "Somersault tightness"
- **Description:** "We found that the tightness of your __*insert position of the somersault*__ was __*insert average hip bend angle during the somersault phase*__ degrees on average. This is rated as __*insert somersault tightness percentile score*__ percentile. Here are some examples of your position in the somersault."
- **Visuals:** GIF of the diver in the somersault phase of the dive
- **Score:** somersault tightness percentile score scaled to be out of 10 points

6.2.5 Knee Straightness Error (not applicable if the position of the dive is tuck, in which case will not be included in the report)

- **Error:** "Knee Straightness"
- **Description:** "We found that your knees bent __*insert average knee bend angle*__ degrees on average. This is rated as __*insert knee straightness percentile score*__ percentile."
- **Visuals:** shares the same GIF as for somersault tightness, which shows the somersault phase of the dive
- **Score:** knee straightness percentile score scaled out of 10 points

6.2.6 Twist Straightness Error (not applicable if no twist phase is detected in the dive)

- **Error:** "Twist Straightness"
- **Description:** "We found that the tightness of your twist was __*insert average hip bend angle during twist phase*__ degrees on average. This is rated as __*insert twist straightness percentile score*__ percentile. Here are some examples of your position in the twist."
- **Visuals:** GIF of the diver in the twist phase of the dive
- **Score:** twist straightness percentile score scaled out of 10 points

6.2.7 Verticalness (over/under-rotation) Error

- **Error:** "Verticalness (over/under rotation)"
- **Description:** "We found that you deviated from vertical by __*insert angle at which the diver deviated from vertical*__ degrees, which was the __*insert verticalness percentile score*__ percentile."
- **Visuals:** GIF showing the diver just before they hit the water during entry.
- **Score:** verticalness percentile score scaled out of 10 points

6.2.8 Body Straightness During Entry Error

- **Error:** "Body straightness during entry"
- **Description:** "The straightness of your body during entry deviated by __*insert hip bend angle at entry*__ degrees, which was the __*insert straightness during entry percentile score*__ percentile."
- **Visuals:** Same GIF as for verticalness, showing the few frames before the diver hits the water.
- **Score:** straightness during entry percentile score scaled out of 10 points

6.2.9 Splash Size Error

- **Error:** "Splash"
- **Description:** "Your splash was __ *insert 'small' if above 50%tile or 'a bit on the larger side' if below 50%tile*__ and was rated at __*insert splash percentile score*__ percentile."
- **Visuals:** GIF showing where splash was detected in the dive.
- **Score:** splash percentile score scaled out of 10 points.

7. Full-size model figure from main paper

Presented in Figure 3.

8. Full Expert Opinions

8.1. Coach A.

We showed Coach A samples of our generated detailed reports and asked for their opinions and feedback. Coach A emphasized the role that overall impression plays in the current judging system: "*Overall impression is very important when I score a dive, and that seems to be overpowered by the entry (especially splash) when humans are judging. I agree that, in practice, we need to slow down and remember previous (non-entry) phases of the dive. The previous phases should matter more when scoring than what current judges do in practice.*" According to Coach A, human judges emphasize overall impression because they cannot always see all of the diver's mistakes in the short amount of time watching the dive so as to come up with a score based on specific errors. Even if judges try to be fair to all aspects and phases of a dive, as humans, it is almost impossible to see the dive perfectly at all times. Moreover, judging based on overall impression is inherently subjective and vulnerable to bias. This kind of unintentional bias shows itself in many ways. For example, if a diver comes into the competition with a strong reputation, judges may give them the benefit of the doubt and penalize mistakes less harshly. Another example is when a diver nails almost all the aspects of the dive (has a great start, jumps high, has a tight tuck, has a strong and confident come-out[1]), but maybe comes out a bit early and lands under-rotated in the water. These dives are often scored very low (3.5-4/10) even when the timing of their come-out was the only mistake they made, showing how heavily weighted the entry is on the overall score. Coach A believes that with the support of our NS-AQA system, scoring in diving can be "*more precise and have less errors.*" Coach A believes that this system, with its generated detailed report, would be helpful in a number of ways. In particular, it would be useful as a tool to (1) teach judges how to score, (2) catch any errors that a human judge may miss, (3) settle disagreements between judges, (4) encourage safety by penalizing dangerous actions (e.g. getting too close to hitting the board) that lots of human judges overlook.

8.2. Judge B.

We also showed our AQA results to a certified judge on the Judges Commission, Judge B. Judge B shared their experiences as a diving judge for over 30 years. They envisioned our system as an educational tool for teaching judges, coaches, and divers

[1]A diver's "come-out" refers to when they straighten their body out to slow the somersault in preparation for entry. This should ideally be one quick, dynamic motion.

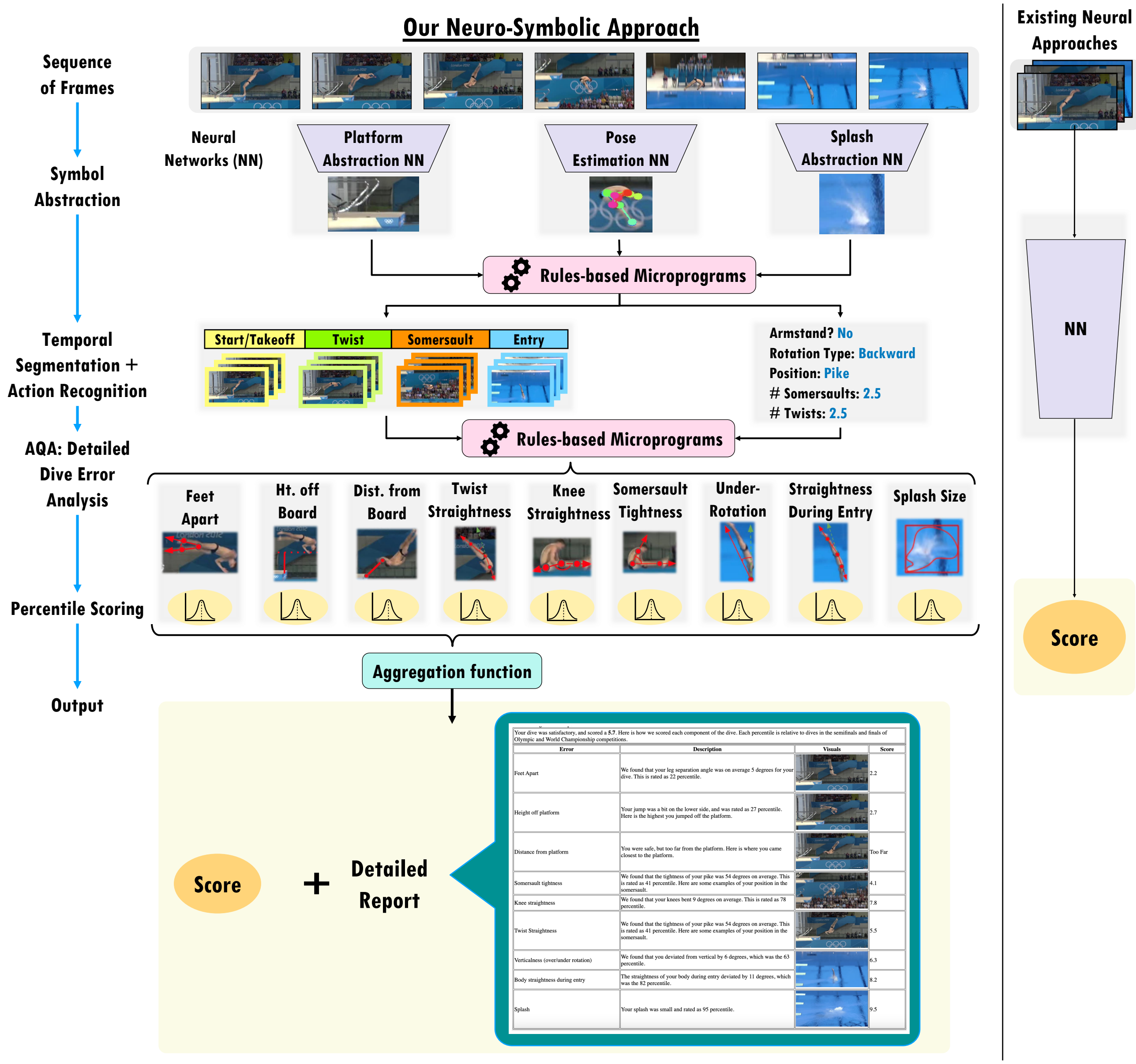


Figure 3. **Neuro-Symbolic Action Quality Assessment (NS-AQA) vs Neural AQA.** Our NS-AQA approach (Left) employs neural networks to extract crucial symbolic information, such as platform location, frame-by-frame pose estimation, and splash detection. These symbols furnish objective data utilized for rules-based fine-grained action recognition, temporal segmentation, and detailed error analysis. The outcome is an objective score and a comprehensive visio-linguistic report, complete with supporting visual evidence, generated programmatically. This is much more valuable than existing AQA approaches (Right) that can only predict a single score (potentially biased) without any accompanying explanation.

how to break down the dive into all of its components. "*We're humans, we can't get it all right,*" Judge B says, especially since judges only get 3 seconds to score a dive. Judge B particularly emphasized our system's potential to improve diving safety due to its automated "distance from board" measurement. They told an anecdotal story that happened during a Senior National Competition in which they were judging. One of the divers "*just had too tight of an inward 3.5*," meaning he was just too close to the platform to be safe. In diving, hitting the platform can be the difference between life and death, so it is very important to prioritize the safety of the diver over anything, even if that means scoring their dive very low. Judge B told us that they raised their hand to say that the dive was too close, which lowers the maximum score a judge can give from 10 to 2. Judge B received a lot of backlash from the diver's coaches, teammates, friends and family because they took away his spot at the World Championships with that call. "*There aren't many judges with the guts to make a call like that, especially thinking about how hard the diver has worked for that very moment, just for a judge to say they were too close to the board,*" Judge B tells us. "*But if we don't keep it safe, we don't have a sport.*" Judge B is "*very excited for this system to be able to make that call for us*," as it moves the *villain*-role from the human judges to the AI to make those difficult decisions.

References

[1] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. 1

[2] Eric Goodman. Diving 101: Scoring. 6

[3] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. 1

[4] Mahdiar Nekoui, Fidel Omar Tito Cruz, and Li Cheng. Eagle-eye: Extreme-pose action grader using detail bird's-eye view. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 394–402, 2021. 1

[5] Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang. Deep high-resolution representation learning for human pose estimation, 2019. 1

[6] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. `https://github.com/facebookresearch/detectron2`, 2019. 1

[7] Jinglin Xu, Yongming Rao, Xumin Yu, Guangyi Chen, Jie Zhou, and Jiwen Lu. Finediving: A fine-grained dataset for procedure-aware action quality assessment, 2022. 1