# Supplementary - Moment of Truth: Dealing with Negative Queries in Video Moment Retrieval

Kevin FlanaganDima DamenMichael WrayUniversity of BristolUniversity of BristolUniversity of Bristolkevin.flanagan@bristol.ac.ukdima.damen@bristol.ac.ukmichael.wray@bristol.ac.uk

In this supplementary, we provide more information about the dataset creation in Sec. 1, describing the process of generating the out-of-domain negative queries using LLMs and demonstrating our prompts and categories as well as details about the Negative-Aware Video Moment Retrieval dataset. We display full results for SVMs trained on output saliency scores for all three models (UniVTG [3], CG-DETR [4], QD-DETR [5]) in Sec. 2. We expand on the adjustments made to the losses for UniVTG-NA and detail the QD-DETR and CG-DETR implementations in Sec. 3. We demonstrate the out-of-domain generalisability of UniVTG-NA in Sec. 4. We further motivate the need for negative-aware methods in video moment retrieval by displaying results with negative queries from UniMD [6] in Sec. 5. Finally, we show more qualitative results from UniVTG-NA on QVHighights and Charades-STA in Sec. 6.

#### **1. Dataset Information**

#### 1.1. Out-of-Domain Negative Query Generation

As mentioned in the main paper, the out-of-domain negative query sentences were generated using a large language model (LLM). Four broad scenarios were used as query topics, these were "competitive sport", "animal behaviour", "physics laboratory", and "mathematics class". In Table 1, the subtopics for each of these topics are listed. The prompt and specific LLM used for each topic is also displayed. Prompts were empirically chosen to ensure the quality and diversity of the generated sentences. For example, for the "animal behaviour" topic, it was found that using scientific names improved upon these aspects, hence most of the subtopics are scientific names. The four scenarios were chosen as they represent scenarios which are unlikely to be present within the OVHighlights and Charades-STA datasets, which cover news, vlogs and household actions. The choice of 4 broad scenarios helps to ensure that the OOD negatives remain OOD and do not accidentally produce false negatives. By using prompts specifically describing the actions as short, unique and varied, we are able to get a wider range of sentences without the language becoming too decorative. Sample sentences from each scenario are displayed in Table 2. We use the same set of OOD Negatives for both QVHighlights [2] and Charades-STA [1].

#### **1.2. Negative Aware Dataset Details**

Table 3 displays the number of positive and negative queries used during training/evaluation of the models. For Charades-STA, where there are fewer negative queries than positive for out-of-domain, the negative queries are assigned to multiple videos. This still produces a distinct signal as each video-sentence pair offers a different semantic relationship.

#### 2. SVM Trained on Saliency Scores

We train an SVM on the outputted positive and negative query saliency scores from UniVTG [3], QD-DETR [5] and CG-DETR [4] for the QVHighlights and Charades-STA datasets. This is to quantify how separable positive and negative queries are when the relationship between them is modelled using saliency outputs from the base models, without any explicit training for negative rejection. Results are displayed in Table 4.

The SVM results on QVHighlights show high rejection accuracy at the cost of decreased R1@ $\theta$  scores. In the case of QD-DETR, these are significantly decreased. The Charades-STA results show reasonable rejection accuracy at significant cost to the R1@ $\theta$  scores for CG-DETR and QD-DETR, while UniVTG fails to achieve high rejection accuracy but has better R1@ $\theta$  scores. Overall these results display the limitations of using the saliency outputs from the base models alone for combined moment retrieval and negative rejection, particularly on datasets without ground-truth saliency scores such as Charades-STA. It further motivates the need to train explicitly for negative rejection.

Topic	Competitive Sport		Animal Behaviour			Physics Laboratory	Mathematics Class
Model	Chat-GPT (GPT-40)		Claude 3 Opus			Claude 3 Opus	Claude 3 Opus
Prompt	Generate X sentences describing actions in <subtopic></subtopic>		Generate X unique and varied short sentences of visual actions carried out by <subtopic></subtopic>			Generate X unique and varied sentences of visual actions carried out by a person working in a <subtopic> lab.</subtopic>	Generate X unique and varied sentences of visual actions carried out by a person working in a <subtopic> class.</subtopic>
	athletics field events baseball cricket darts field hockey gymnastics	badminton boxing cycling fencing golf ice hockey	anatidae anura bovidae cephalopods chiroptera crocodilia	anguilliformes big cats camelid cervidae chondrichthyes decapods	annelids bivalves canidae chelicerata cnidaria echinoderms	atmospheric physics biophysics chemical physics classical mechanics condensed matter physics cosmology	applied mathematics calculus combinatorics computational maths geometry graph theory
	lacrosse rugby skateboarding snooker soccer swimming tennis water polo	rowing running skiing snowboarding squash table tennis ultimate frisbee	hymenoptera lepidoptera monotremes pinnipeds primates rodents stomatopods talpidae ursidae	gastropous insects lizards mustelids platyhelminthes proboscidea serpentes strigiformes testudines wading birds	lagomorphs marsupials osteichthyes porifera ratites spheniscidae suina urodela	electroniaghetishi electronics fluid dynamics geophysics medical physics optical physics particle physics quantum mechanics thermodynamics	probability statistics

Table 1. List of topics and subtopics used for out-of-domain negative generation, along with the prompts and LLMs used. X represents the number of sentences requested which varied from 50 to 100.

## **3. Model Details**

## 3.1. UniVTG-NA

For UniVTG-NA, the input to the classification head is a direct sum of the indicator scores and saliency scores. *i.e.*  $g_i = f_i + s_i$  where  $g_i$  is the classification head input at index *i*.

**Loss Adaptations.** We specify the adjustments made to the losses for the UniVTG-NA model from UniVTG [3]. Aside from the boundary prediction losses being set to 0 for the negative queries, the saliency losses are also adjusted. UniVTG uses a saliency loss  $\mathcal{L}_s$  which is a weighted summation of inter-video and intra-video contrastive losses. It is not possible to use the contrastive saliency loss with negative queries. Therefore, for negative queries the saliency loss is defined as a loss applied directly on the cosine similarity between the video clip  $\mathbf{v}_i$ , and sentence features  $\mathbf{S}$ , with  $\lambda_s^-$  as the loss weighting.

$$\mathcal{L}_{s}^{-} = \lambda_{s}^{-} \cos(\mathbf{v}_{i}, \mathbf{S}) := \lambda_{s}^{-} \frac{\mathbf{v}_{i}^{T} \mathbf{S}}{\|\mathbf{v}_{i}\|_{2} \|\mathbf{S}\|_{2}}$$
(1)

This is done as the saliency scores are computed via cosine similarity between sentence and video clip features for Uni-VTG, so achieves our principle of designing the saliency loss for negatives such that it pushes the saliency scores lower. Furthermore, for UniVTG-NA's foreground matching loss with negative queries,  $\mathcal{L}_f^- = \mathcal{L}_f$  as no adjustments are made to the matching loss, which is a BCE loss on the individual indicator scores. This already achieves the aim of pushing the indicator scores lower.

## 3.2. QD-DETR-NA & CG-DETR

Negative-aware versions of QD-DETR and CG-DETR were also trained to evaluate the proposed method on other models. The details of the QD-DETR and CG-DETR implementation are as follows: Given the indicator score outputs  $\{\tilde{f}_1, ..., \tilde{f}_M\}$  and saliency score outputs  $\{\tilde{s}_1, ..., \tilde{s}_{L_v}\}$ , the input to the classification head is a concatenation  $g = \{\tilde{s}_1, ..., \tilde{s}_{L_v}, \tilde{f}_1, ..., \tilde{f}_M\} \in \mathbb{R}^{(L_v+M)}$ , where  $L_v$  is the number of video clip features and M is the number of moment queries. This implementation is represented in Figure 1. This is chosen as opposed to a summation because QD-DETR/CG-DETR use moment queries to generate the moment candidates rather than just the text-attended video clip representations from the encoder. In this case, there is not a one-to-one correspondence with the saliency scores, *i.e.*  $L_v \neq M$ .

As with UniVTG, the boundary losses were set to 0 and the foreground matching loss was retained for the negative queries. For both methods, the saliency loss has three components, two of which are contrastive and are therefore not feasible for negative queries. The remaining loss works to reduce the negative query saliency scores, thus achieving the principle aim of the negative query saliency loss. Therefore it is retained as the sole loss for negative queries. It is shown for a saliency score output  $s_i$  with loss weighting  $\lambda_s^$ below.

$$\mathcal{L}_s^- = \lambda_s^- (-\log(1 - s_i)) \tag{2}$$

The outlook theore is the series of	Competitive Sport	Animal Behaviour
The opposite the shore hole does by a lob posite prior hole does not prior to a streng when a reading streng the streng	The outfielder throws to home plate	An osprey dives into the water, snatching a fish with its talons
The share head are he	The opponent hits a drop shot followed by a lob	A northern harrier glides low over a meadow, searching for small mammals.
The period of possing term derived in their own entrony will a vell-back tack. A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected bill derived term of a first over and African granukal A selected were many term of a first over and African granukal A selected were many term of a first over and the over and t	The striker heads the ball into the net	An ovambo sparrowhawk sits near its nest, guarding its eggs.
The place reacts a deceptic stack hand dep and of the stack of the sta	The punter pins the opposing team deep in their own territory with a well-placed kick	A slender-billed kite hunts for insects over an African grassland
The opport affects a startistics drop into the kackom         A lamproy own in a signed equip drafts. Acting theremuse truth for potential mass.           The opport actions to opport any facts into the startistic draft action into the startis draft action into the startistic draft action into the starti	The player executes a deceptive backhand drop shot	A white-backed vulture strips meat from a carcass with its strong beak
The goad of largery data served lay set?       A goad of largery data set of the data	The opponent flicks a shuttlecock deep into the backcourt	A lamprey swam in a figure-eight pattern, leaving pheromone trails for potential mates
The place without is a concert built of the same of the scale of	The goalie makes a sprawling save	A group of lamprey larvae anchored themselves to rocks, facing into the current
The opport over any aboding the boos         A topace diated integring in action of cyclic wite, ite, good book minimum failer code           The order over event on the over over any aboding in action of cyclic wite in a first integring to a code of cyclic wite integring in a code of cyclic wite integring	The player switches to a colored ball after potting all reds	A bronze eel lay coiled on the seafloor, its coppery scales gleaming
lie carf michael he hall o is carmine bedre here jushed A bodre were minimie here is sound a prec of dimitsed Lie vere minimie indiver intege inte	The opponent covers up, absorbing the blows	A topaz eel darted through a school of yellow tang, its golden body mirroring their color
Ide black during the spong large is and with it drops the spong large is a lard with it drops the spong large is a lard with it drops the spong large is a lard with it drops the spong large is a lard with it drops the spong large is a lard with it drops the spong large is a large in large is a large is large is a large is a large is a large is larg	The center offloads the ball to a teammate before being tackled	A bootlace worm tangles itself around a piece of driftwood
in the overage of the stand back text is also due to the stand	The flanker disrupts the opposing team's maul, forcing a turnover	A red-eyed tree frog clings to a lead with its sticky toe pads
In the statum to the dock at with intervalue, the statu intervalue, the statu intervalue, the statu intervalue i	The rowers maintain their balance as the boat rocks gently on the water	The Amazon milk frog inflated its body, trying to appear larger
In each mathematication of the stand of	The rowers return to the dock and disembark from the boat	The jaguar's powerful jaws crushed the turtle's shell
In the action of a constraint o	The archer lands of dashe series with end with end of the larger	A long chain extends its signoins, drawing in water to inter out rood particles
Intermediation and the scale on the scale of the Scale on the scale of the Scale on th	The scatter randos a double axet with precision	A kulu reached up to provise on acacta tree reaves
The called calls of the max on the coder view of the bar         The codes ployses don'the tander plots of the store ployses of the tander plots of the store plots and backgroup of a lower           The calls explored backgroup of the bar         Anonoral's large, consule) ploys resemble a cluster of full moos.           The resemble with precision on the ord exvices material         The coreacy interaction of the store ployses on the store in the store of full moos.           The resemble most of the store of the store exosts material         The coreacy interaction of the store of the st	The swimmer's streamine position reduces resistance infogn the water	A case for the playturity wresties with its storing outside their den
The stare recents a pump continuition, linking import officeren rotations.         A can be over the rotation set the stare in the set of the rotation of of the rota	The stillate lends on the must on the state side of the her	A cape tox, known for its incomma nabits, energies norm is used at dusk to begin minimig
The exercise sectors and space or sectors and space	The durice range of the first of the other size of the ball	A arek nidor amidra and ha form it hiding and in a dawar
Interpretation         Procession         Mathematics (Class)           Interpretation         Mathematics (Class)           The secontice measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the sound absorption coefficient of the level assorption measured the level assorption measured the level assorption measured the level assorption measurem the level assorption the level as a set level assorption the level assorption the level assorption the level as a set level assorption the level assorption the level as a set level assorption the level as a set level assorption the level as a set level assorption the level assorptin the level assorption the level assorption the level	The scale executes a jump combination, mixing jumps of unrelease for formers	A mean exercise transfer reserved a beer from its meaning spot in a nower
The reserver measured he sound shorpion coefficient of the sew assound material measure the sound strong per dest affive set sound strong per dest affire set sound s	The gynnast tunbles with precision on the noor exercise mat	A moon cotar starge, founded polyps resemble a cluster of fun moons
The constraint measure of the sound relation is under sound space atom         Fe draws a hore and whilker plot to compare the distribution         different dua sets           The secand set milling plot to compare the sound power of height of the cloud base         For event we come a columne the sound power of the let main of the sound sets of the her possibility of the loud base         For event we columne the probability of the loud base         For event we main and the probability of the loud base         For event we main and the probability of the loud base         For event we main and the probability of the loud base         For event we main and the probability of the loud base         For event we main and the probability of the loud base         For event we main and the probability of the loud base of the lou	The recorded and abcomtion coefficient of the new countin meterial	They exists a flow diagram to show the store in the algorithm
The seried a sound instancia probe to reason the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the series of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation of the sound part of the jer englane mode generation part of the jer englane mode generation of the sound part of the jer englane mode generation part of the jer englane generatio	The reserved the sound absolution coefficient of the new acoustic material	He draws a how and which a lat to commerce the distributions of different date sets
The reactive number of the chemic area of the chemic area of the sequencing of variables to change us of the the partial differential quegation of the chemic area of the sequencing of variables to change us of the the partial differential quegation with experiment of the sequencing of variables to change us of the the partial differential quegation with experiment of the chemic area of the sequencing of variables to change us of the theorem of the sequencing of variables to change us of the variables of the sequencing of variables to change us of the variables of the variables to change us of the variables of the variable of the variable of the variable of the variable of the variables of the variable of the varia	The accustician measure in evolution mack of the window using a pink holse generator	They are a Van diagram to find the probability of the union of the entry and sets
With stack that the cheating under a capillar puthe to load the viscous souic liquid into the cheating as the product status of the cheating status of t	The researcher uses a callometer to determine the height of the cloud base	She uses the separation of variables technique to solve the partial differential equation
The granue student intent y studies the XPS spectrum, identifying the chemical states of the elements present on the calalyst strucc Be carefull by obtioned by sample in the color of the phit-coli magnet The scientist replaced the filament in the electron gun Fe carefull positioned the sample in the calaryst strucc He carefull positioned the sample in the calaryst strucc The carefull positioned the sample in the calaryst strucc The scientist adjusted the struccu struc	With a steady hand the chemist uses a capillary tube to load the viscous ionic liquid into the rheometer for flow behavior studies	He arranges a set of numbered tiles to illustrate the concent of permutations with repetition
See records the funk room the oscilloscope in her lab rootbod         He labble such vertex with a unique letter, making it exist to rief to opecific nodes           The sciential replaced the filament in the electron gamet         He labble and here to with a unique letter, making it exist to rief to opecific nodes           The sciential replaced the filament in the electron gamet         He draws a graph with a uninium, spanning tres, a subgraph connecting all vertices with the minium total edge weight           The sciential replaced the signest roots medium using magnetic roosance (many signest roots mediam using mathematical echicage         The statistical needed to large printed graph, ensuing the connection between combinators of norbitation of the connection pagnetic roosance (many signest roots medium using magnetic roosance (many signest roots medium using mathematical echicage           The statistical roots and biomitatical the fore printed graph of the light sign a spectrometer         He eators and its conduction in the fore root operation of a norbitatical the fore printed graph of printed magnetic roosance (many signes danorots medium using mathematical echicage) </td <td>while a second plant, the electricit does a capitally tube to load the viscous lotter induct into the moment on how obtained states of the alements present on the existing states of the alements present on the exi</td> <td>With a critical was sha argument the nartial demendance plots assessing the impact of individual features on the model</td>	while a second plant, the electricit does a capitally tube to load the viscous lotter induct into the moment on how obtained states of the alements present on the existing states of the alements present on the exi	With a critical was sha argument the nartial demendance plots assessing the impact of individual features on the model
The scientist pipeded the future in the electror of pike pil-coil mager       She shades a verte to indicate it has ben visited during a graph traversal         The scientist pipoted the future of the feettor og pike       She shades a verte to indicate it has ben visited during a graph traversal         The scientist pipoted the future of the spector og pike pipoted is posted to pipote pipote pipoted is the pipoted is a pipoted pipoted is pipoted pi	The product a from the oscilloscope in her lab notebook.	He labels each vertex with a unique letter making it easier to refer to specific nodes
He cardially passioned the sample in the care of no-split-coll magnet       He draws a graph with a minimum spanning tree, a subgraph cancucting all vertices with the minimum tool edge weight         The scientist diguided the string on the surface plasmon resonance (RPM) instrument       He draws a graph with a minimum spanning tree, a subgraph cancucting all vertices with the minimum tool edge weight         The scientist diguided the string on the surface plasmon resonance (RPM) is sufficient on the partner is a provide in the diguided the large print edge in the diguided the partner in the diguided the align on the cancen to engline on the diguided the partner in the diguided the diguided the partner in the diguided the partner in the diguided the diguided the partner in the diguided the dinter diguided the dinted diguided the dinter diguided the diguide	The scientist replaced the filament in the electron sun	She shades a vertex to indicate it has been visited during a graph traversal
The scrinit adjusted the settings on the surface plasmon resonance (RPR) instrument       He shades the frae representing the union of two probability events         The consoligits cardfully positioned the spectrography the light from a distant supernova       He shades the frae representing the union of two probability events         The scander cardfully folded the spectrography the light from a distant supernova       He shades the frae representing the union of two probability events         The scander cardfull folded the spectrography the same cardfully folded the large privated graph, ensaring the crasses were sharp and the edges aligned         The scander cardfull folded the spectrography the same cardfull folded the large privated graph, ensaring the crasses were sharp and the edges aligned the cardfully folded the rank or execut at exit's weares.         The scander cardfull folded the spectrography the same cardfull folded the large privated graph, ensaring the chalbback representing the union of two presented to divert and the edges aligned the sample same cardfull folded the large privated graph, ensaring the chalbback representing the cardfull same cardfull folded the large privated graph, ensaring the chalbback representing the endets of the same endets aligned the same for and the edges aligned the number of combinations and the edges aligned the number of combinations and the edges aligned the adjacency relations is a normalization and the edges aligned the adjacency relations is a combinations and bindit exerce endets aligned the adjacency relation shifts and combinations and bindit exerce and the edges aligned the configurations, exploring the number of combinations and bindit exerce and the edges aligned the configurations.         She dinstant supernova distant supernova electro	He carefully positioned the sample in the center of the split-coil magnet	He draws a grant with a minimum spanning tree a subgrant connecting all vertices with the minimum total edge weight
The cosmologist carefully positioned the spectrograph, ready to analyze the light from a distant supernova       The statistician carefully folde the large printed graph, ensating the creases were sharp and the edges aligned         The statistician carefully folde the large printed graph, ensating angenice resonance imaging       The statistician used a highly liter to trace the read in lice on the time series plot         The graphysicist calibrate the radionuclide calibrate for accurate activity measurements of radiopharmaceuticals       The statistician used a highly liter to trace the read use of a new mathematical technique.         The scientist adjusted the position of the camere to capture the desire divergence activity measurements of radiopharmaceuticals       The mathematician creates a Pascal's triangle, highly lighting the connection between combinatoris and binomial coefficients.         She calculate the probability of a defective product using quality control dan       She writes out the formal in for accurate activity the material         She calculate the probability of a defective product using quality control dan       Materian actual divergence and the deges aligned         The accordinate's fingers danced across the keys, effortIssly transitioning between notes       Materian actual divergence and the deges aligned         The accordinate's fingers danced across the keys, effortIssly transitioning between notes       Materian actual transition actual transition actual transition actual transition active production activity of a sub transition activity of a sub transitity of a sub tradia divergence and transition activity of	The scientist adjusted the settings on the surface plasmon resonance (SPR) instrument	He shades the area representing the union of two probability events
He studies the flow patterns in a porous medium using magnetic resonance imaging       The analyst used a highlighter to trace the trend line on the time series plot         The researcher measures the hermal conductivity of a rock sungue sing a divided by appratuss       The analyst used a highlighter to trace the trend line on the calibboard, representing the regression equation         The psychicit aties is schmidt hammer to test the strength of a rock sunce on the analyst or a capture the discing integression equation       The statisticatica used a chilk line to dwas a perfectly stright line on the chilkboard, representing the regression equation         The psychicit aties is discing on the camera to capture the discing imaging control dat       The mathematican creates a pacel's triangle, highlighter to consochion between combinatoris: and biomial coefficients         She calculates the probability of a defective product using quality control dat       He creates a matrix to represent the alignecary relationships in a combinatorial graph         She calculates the probability of a defective product using quality control dat       He arranges a set of dominose infigerent conjugations, exploring the number of possible tillings         She accordinis's fingers danced across the keys, effortlessly transitioning between notes       Musician Performance         The baccordinis's fingers danced across the keys, effortlessly transitioning between notes       Musician Performance         The baccordinis's fingers danced across the keys, effortlessly transitioning between notes       He arranges a set of dominos on a negressitic second second seterify setand second second second seterify setand seco	The cosmologist carefully positioned the spectrograph, ready to analyze the light from a distant supernova	The statistician carefully folded the large printed graph, ensuring the creases were sharp and the edges aligned
The searcher measures the thermal conductivity of a rock sample using a divided bar apparatus       The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation         The physicist calibrated the radionaclide calibration for accurate activity measurements of radiopharmaceuticals       The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation         The physicist calibrated the radionaclide calibration for accurate activity measurements of radiopharmaceuticals       The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation         The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation       He leans forward, listening intently to his colleague's explanation of a now mathematical chanya         The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation       He intently chalkboard, representing the regression equation         The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation       He intently chalkboard, representing the regression equation         The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation       He intendly chalkboard, representing the regression equation         The statistican used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation       He intendly chalkboard, representing the regression equation	He studies the flow natterns in a norous medium using magnetic resonance imaging	The analyst used a highlighter to trace the trend line on the time series plot
The geophysicit uses a Schmidt hammer to tes' the strength of a rock ouroop in the physicia clicharadd the radionaclide calibration for accurate activity measurements of radiopharmaceuticals the physicia clicharadd the radionaclis of a schwart activity measurements of radiopharmaceuticals the physicia clicharadd the radionacles a Pascal' thradgel, highlighting the connection to setteme the desired image in the distingt of the calculation of the camera to capture the desired image in the distingt of the calculation of the camera to capture the desired image in the distingt of the dist	The researcher measures the thermal conductivity of a rock sample using a divided bar apparatus	The statistician used a chalk line to draw a perfectly straight line on the chalkboard, representing the regression equation
The physicist calibrated the radionuclide calibrator for accurate activity measurements of radiopharmaceuticals       The mathematician create's a Pascal''s triangle, highlighting the connection between combinatories and binomial coefficients         The scientist adjusted the position of the camera to capture the desired image       She writes out the formula for calculating the number of consibility on so on objects take n r a t a time         She alculates the probability of a defective product using quality control data       He creates a matrix to represent the adjacency relationships in a combinatorial graph         She alculates the probability of a defective product using quality control data       He mathematician draws a tree diagram to illustate the Collar conjecture         She adjuste the positive for control the interference between the microwave signals       The mathematician draws a tree diagram to illustate the Collar conjecture         The mathematicin draws a tree diagram to illustate the Collar conjecture       He mathematician draws a tree diagram to illustate the Collar conjecture         The mathematician draws a tree diagram to illustate the Collar conjecture       He mathematician conversite in the Collar conjecture         The mathematicin draws a tree diagram to illustate the Collar conjecture       He mathematician draws a tree diagram to illustate the Collar conjecture         The mathematicin draws a tree diagram to illustate the Collar conjecture       He mathematician draws a tree diagram to illustate the Collar conjecture         The mathematician draws a tree diagram to illustate frequence       He mathematician draws a tree dia	The geophysicist uses a Schmidt hammer to test the strength of a rock outcrop	He leans forward, listening intently to his colleague's explanation of a new mathematical technique
The scientist adjusted the position of the camen to capture the desired image       She writes out the formula for calculating the number of combinations of n objects taken r at a time         He measured the wavelength of the light using a spectrometer       He creates a matrix to represent the adjacency relationships in a combinational graph         She adjusted the probability of a defective product using quality control data       He creates a matrix to represent the adjacency relationships in a combinational graph         She adjusted the phase shifter to control the interference between the microwave signals       The mathematicina draws a tree diagram to illustrate the Collar conjecture         The scientific uses a laser themometer to measure the surface temperature of the material       Mewrites out a proof using mathematical induction, establishing a pattern         The basic player's hands alterned earons the keys, effortlessly transitional pettween notes       He material induction, establishing a pattern         The cellst learnet, convery the temotion of the picce through their posture       He ranses a set of dominose in different configureant         The cellst learnet, convery the temotion of the picce through their posture       He ranses as lear the moment, convery the temotion of the picce through their posture         The class learnet to measure the use adjusted of the stage light as she played a upber time.       He ranses as learnet the material         She brites the strings on the fricture configureant to string a soc.       He ranses as learnet the material induction.         She brites the string sof the probl	The physicist calibrated the radionuclide calibrator for accurate activity measurements of radiopharmaceuticals	The mathematician creates a Pascal's triangle, highlighting the connection between combinatorics and binomial coefficients
He neasured the wavelength of the light using a spectrometer       He creates a matrix to represent the algencery relationships in a combinatorial graph         Be cadulates the probability of a defective product using quilty control dan       He creates a matrix to represent the algencery relationships in a combinatorial graph         Be adjusts the phase shifter to control the interference between the microwave signals       He arranges as at of domineos in different configurations, exploring the number of possible tilings         The active states also the the probability of the number of the material       Musician Perturbations       He arranges as at of domineos in different configurations, exploring the number of possible tilings         The active states also the biothypic to fill the bag with air       He arranges as at of domineos in different configurations, exploring the number of possible tilings         The mathematician finance ting espliciting pattern       He arranges as at of the state actore offect       He match matrix to represent the algencery relationships in a gombinatorial graph         The bag player's hands alternate between drums       He creates and the number of the pice through their posture       He reatents       He reatent the state state actore offect         The bag player's hands alternate between drums       He creatent as num's to represent the algencery relationships in a gombinatorial graph       He creatent as num's to represent the algencery relationships in a combinatorial graph         The target as at off the state information of the pice through their posture       He arranges as at ot pice the state inform	The scientist adjusted the position of the camera to capture the desired image	She writes out the formula for calculating the number of combinations of n objects taken r at a time
She calculates the probability of a defective product using quality control data       He arranges a set of dominoes in different configurations, exploring the number of possible tillings         She adjusts the phase shifter to control the interference between the microwave signals       The mathematician draws a tree diagram to illustrate the Collatz conjecture         The scientist uses a laser themmenter to measure the surface temperature of the material       Musician Performance         Musician Performance       Musician Performance         The anabows into the blowpips to fill the bag with air       He arranges a set of dominoes in different configurations, exploring the number of possible tillings         The basic player's hands moved in a blur, creating a structure fingerprixing pattern       He writes out a poof using mathematical induction, establishing a pattern         He eurole strings vish his palm, creating a structure fingerprixing pattern       He writes out a poof using mathematical finderprixing pattern         He cell steans into the instrument, conveying the emotion of the picce through their posture       He arranges on the friedword (creating a soft, sizzling soud)         The interfest power during of out as they blow into the motipicce       He arranges on the friedword, creating a dynamic sound         She brushes the structure during on a structure of the stap influx could a sound blay of power during a drawing on the hands on the kary is the material       He arranges on the friedword (creating a soft, sizzling a soft, sizzling a soft during the prefest on the pedia and her hands on the kary is the prefers       He arranges	He measured the wavelength of the light using a spectrometer	He creates a matrix to represent the adjacency relationships in a combinatorial graph
She adjusts the phase shifter to control the interference between the microwave signals The mathematician draws a tree diagram to illustrate the Collarz conjecture   The scientist uses a laser thermometer to measure the surface temperature of the material She writes out a proof using mathematical induction, establishing a pattern   The accordionist's fingers danced across the keys, effortlessly transitioning between notes Musician Performant   The math hows into the blowpipe to fill the bag with air Musician Performant   The how into the blowpipe to fill the bag with air He mutel the strings with his plant, creating a intricate fingerpicking pattern   He mutel the strings with his plant, creating as conto offect He mutel the strings with his plant, creating a string sound   The be cells telanes, converying the motion of the picce through their posture He mutel the strings of the string sound is plant.   The inclusion of the string sound is plant, creating a percussive rhythm He advermates between blown into the moutpicce   She plants the string of out a string to a stool He advermate between the harmonica, creating a dynamic sound   With closed eys, the musician swayed gently as the string blog data strings He advermate between blown gan drawing on the harmo's delicate strings   She gently presses the white keys with the fingerpick He mutel the strings with is planterning   She sinde is his first hands on the keys She forthead and the hardo is the string in the advermate between the sheer blow in the material advermate in the interformate between the sheer blow in the material advermate in the interformate between the sheer blow in the hardo ist	She calculates the probability of a defective product using quality control data	He arranges a set of dominoes in different configurations, exploring the number of possible tilings
The scientist uses a laser thermometer to measure the surface temperature of the material       She writes out a proof using mathematical induction, establishing a pattern         Musician Performance         The accordionist's fingers danced across the keys, effortlessly transitioning between notes         The man blows into the blowpipe to fill the bag with air         The bair players' hands moved in a blur, creating a stractate fingerpicking pattern         He muted the strings with his plam, creating a stractate offect         The boogn player's hands alternate between drougs         The cellist lease into the instrument, conveying the emotion of the piece through their posture         She brushes the strate droug a soft, sizzing soud         Their checks puict out as the book into the motipice         She plays the guitar while sitting on a stool         He atagets between blowing and drawing on the harmonica, creating a dynamic sound         With closed eyes, the musican wayed gently as they strummed the harp's delicate strings         She gently presses the white keys with the fingertips         She gently presses the white keys with the fingertips         She since alt the audience, for savophone gleaning under the stape lights as she played a upbeat tune         He sildes his the thands on the keys         She since alt the audience, for savophone gleaning under the stape lights as she played a upbeat tune         He sildes his for thands cords the keys         She prince field	She adjusts the phase shifter to control the interference between the microwave signals	The mathematician draws a tree diagram to illustrate the Collatz conjecture
Musician Performance           Musician Performance           The accordionist's fingers danced across the keys, effortlessly transitioning between notes           The man blows into the blowpipe to fill the bag with air           The man blows into the blowpipe to fill the bag with air           The basic player's hands moved in a blur, creating an intricate fingerpicking pattern           He muted the strings with his playme, creating a staccal offect           The boos player's hands alternate between drums           The cells! teams, converying the emotion of the piece through their posture           She brushes the stare drum lightly, creating a soft, sizzling sound           Their checks put out as they blow into the moutprice           She plays the guitar while sitting on a stool           He atpreade bits fingers on the freboard, creating a dynamic sound           He atpreade bits mayed gettly as they strummed the harp's delicate strings           She genty preses the while keys with her fingerprisp           She mild at the audience, her sucryohne gleaning under the stage lights as she played a upbeat tune           He skides his left hand dong the strings to change the pitch           The studes his der hands steady for a long, sustained note           He stilds his the transhow steady for a long, sustained note           He stilds his the transhome and his fingers, ensuring he played each note correctly	The scientist uses a laser thermometer to measure the surface temperature of the material	She writes out a proof using mathematical induction, establishing a pattern
The accordionist's fingers danced across the keys, effortlessly transitioning between notes The man blows into the blowpip to foil II the sky with air The banjo player's hands atmoved in a blur, creating an intricate fingerpicking pattern He muted the strings with his palm, creating a staccato effect The banjo player's hands atment between drums The ellist leans into the instrument, converying the emotion of the piece through their posture She brusks bet ken stare drum lightly, creating a soft, sizzling sound Their checks puff out as they blow into the moutpiece She plays the guitar while sitting on a stool He atternate between drums, creating a percussive rhythm He atternates between blowing and druming on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they stronger the stage lights as the played a upbeat tune He sides his left hands on the keys She endly presess the white keys such possible stage lights as she played a upbeat tune He sides his left hands on the keys They key their hands steady for a long, sustained note He tilts the trombone up for a high note They key their hands steady for a long, sustained note He tilts the trombone up for a high note	- Musician Perfor	nance
The man blows into the blowspipe to fill the bag with air of the bag with air of the banjo player's hands moved in a blur, creating an intricate fingerpicking pattern He muted the strings with his palm, creating a staccato effect The bongo player's hands alternate between drums The cellist leans into the instrument, converying the emotion of the picce through their posture Bondon The bondon The picce through their posture Bondon The bondon The picce through their posture Bondon The bondon The picce Horough their posture Bondon The picce Horough their posture Bondon The picce Bondon The Bondon The Picce Bondon The Bondon The Picce Bondon The Picce Bondon The Bondon The Picce Bondon The Bondon The Picce Bondon The Bond	The accordionist's fingers danced across the keys, effortlessly transitioning between notes	
The banjo player's hands moved in a blur, creating an intricate fingerpicking pattern He mutch dhe strings with his plant, creating a stocato effect He mutch dhe strings with his plant, creating a stocato effect He collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the emotion of the picce through their posture Be collisi leans into the instrument, conveying the stocating a	The man blows into the blowpipe to fill the bag with air	
He muich the strings with his palm, creating a staccato effect The bongo player's hands alternate between druns The cellist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument, conveying the emotion of the piece through their posture The relist leans into the instrument conveying the emotion of the piece through their posture The alternates between blowing and drawing on the harmonica, creating a dynamic sound He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presses the white keys with her fingerdips She pieces her feet on the pedals and her hands on the keys She striled at the addrene, che reasorybone gleanning under the stage lights as she played a upbeat tune He slides his left hand dong the strings to change the pitch They play a gliasnab by sliding their finger across the keys They play a gliasnab by sliding their finger across the keys They play a gliasnab by sliding their finger across the keys The ther shead one, che reasors the keys The ther there there the shead there, the stage light as a negle each note correctly The was a gliasnab by sliding the first encore the keys The musical is eyes darde between the shead this fingers, ensuring he played each note correctly	The banjo player's hands moved in a blur, creating an intricate fingerpicking pattern	
The bongo player's hands alternate between drums The cellist learns into the instrument, converying the emotion of the piece through their posture The cellist learns into the instrument, converying the emotion of the piece through their posture The cellist learns into the instrument, converying the emotion of the piece through their posture Their cellist learns into the instrument, converying the emotion of the piece through their posture She brushes the snare drum lightly, creating a soft, sizzling sound Their checks puff out as they blow into the mouthpiece She plays the guitar while sitting to a stool He targed his fingers on the fretDoard, creating a percussive drythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presess the while kergy with her finger arcs. She should be strings to change the pitch She should be strings to change the pitch She should be strings to change the pitch He sides his left thand along the strings to change the pitch They play a glissmado by siding their finger arcs so the keys They keep their hands steady for a long, sustained note He tils the trombone up for a high note He musician S yes darded between the sheet thusic and his fingers, ensuring he played each note correctly	He muted the strings with his palm, creating a staccato effect	
The cellis learns into the instrument, conveying the emotion of the piece through their posture She brushes the same drum lightly, creating a soft, sizzling sound Their checks puff out as they blow into the mouthpiece She plays the guitar while sitting on a stool He tapped his fingers on the freborad, creating a precussive thythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician awayed gently as they strummed the harp's delicate strings She gently presses the white keys with her fingertips She gently presses the white keys with her fingertips She gently presses the white keys with her fingertips She smiled at the audience, her saxophone gleanning under the stage lights as she played a upbeat tune He sides his left hand dong the strings to change the pitch They play a glisando by sliding their finger across the keys They play a glisando by sliding heir finger across the keys He tilts the trombone up for a long, sustained note He tilts the trombone up for a high note The musicain 's eyes darded between the sheet music and his fingers, ensuring he played each note correctly	The bongo player's hands alternate between drums	
She brushes the sarae drum lightly, creating a soft, sizzling sound Their checks public win to the mouthprice She plays the guitar while sitting on a stool He tapped his fingers on the freboard, creating a percussive rhythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presess the while keys with her fingerifs She places her feet on the pedals and her hands on the kays She smiled at the audience, her sucception the stage lights as she played a upbeat tune He sildes his left hand dung the strings to change the pitch They play a glissmolb by sliding their finger across the keys They keep their hands steady for a long, sustained note He tils the trombone up for a high note He tils the trombone up for a high note He tils the trombone up for a high note	The cellist leans into the instrument, conveying the emotion of the piece through their posture	
Their checks puff out as they blow into the mouthpice She plays the guitar while sitting on a stool He tapped his fingers on the fretboard, creating a percussive rhythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician awayed genty as they strummed the harp's delicate strings She genty presses the white keys with her fingertips She places her feet on the pedals and her hands on the keys She smiled at the audience, her saxophone gleaming under the stage lights as she played a upbeat tune He slickes his feft hand dong the strings to change the pitch They play ag listando by sliding their finger across the keys They play ag listando by sliding their finger across the keys He tilts the trombone up for a high noe He tilts the trombone up for a high noe	She brushes the snare drum lightly, creating a soft, sizzling sound	
She plays the guitar while sitting on a stool He tapped his fingers on the frebtoard, creating a percussive rhythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed genly as they strummed the harp's delicate strings She genly presses the while keys with her finger the She smiled at the audience, her suscophone gleanning under the stage lights as she played a upbeat tune He slides his left hand along the strings to change the pitch They play ag listando by sliding their finger across the keys They play ag listando by sliding their finger across the keys He tilts the trombone up for a long, sustained note He tilts the trombone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	Their cheeks puff out as they blow into the mouthpiece	
He taped his fingers on the freboard, creating a percussive rhythm He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presses the white keys with her fingertips She places her fet on the pedias and her hands on the keys She places her fet on the pedias and her hands on the keys She similed at the audience, her saxophone gleaming under the stage lights as she played a upbeat tune He slides his left hand along the strings to change the pitch They leave pitch andbe strings to change the pitch They keep their hands steady for a long, sustained note He tilts the troubone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	She plays the guitar while sitting on a stool	
He alternates between blowing and drawing on the harmonica, creating a dynamic sound With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presses the white keys with her fingertips She places her feet on the pedals and her hands on the keys She smiled at the audience, her swaychone gleanning under the stage lights as she played a upbeat tune He sides his left hand dong the strings to change the pitch They play a glismado by sliding their finger across the keys They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musicain's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	He tapped his fingers on the fretboard, creating a percussive rhythm	
With closed eyes, the musician swayed gently as they strummed the harp's delicate strings She gently presess the white keys with her fingerdips She places her feet on the pedals and her hands on the keys She miled at the audience, her successful the stage lights as he played a upbeat tune He slides his left hand long the strings to change the pitch They play a glissmado by sliding their finger across the keys They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musician's seys darded between the sheet music and his fingers, ensuring he played each note correctly	He alternates between blowing and drawing on the harmonica, creating a dynamic sound	
She gently presses the white keys with her fingertips She places her feet on the pedals and her hands on the keys She smiled at the audience, her saxophone gleaming under the stage lights as she played a upbeat tune He sides his feft hand along the strings to change the pitch They play a glissando by sliding their finger across the keys They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musician's eyes dardet between the sheet music and his fingers, ensuring he played each note correctly	With closed eyes, the musician swayed gently as they strummed the harp's delicate strings	
She places her feet on the pedals and her hands on the keys She smiled at the audience, her successform and the stage lights as she played a upbeat tune Her slides his left hand along the strings to change the pitch They play a glissmado by sliding their finger across the keys They keep their hands steady for a long, sustained note He tills the trombone up for a high note The musicing's seys darded between the sheet music and his fingers, ensuring he played each note correctly	She gently presses the white keys with her fingertips	
She smiled at the audience, her saxophone gleaming under the stage lights as she played a upbeat tune He slides his left hand along the strings to change the pitch They play a glissando by sliding their finger across the keys They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	She places her feet on the pedals and her hands on the keys	
He slides his left hand along the strings to change the pitch They play a glissando by sliding their finger across the keys They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musicains' seys darted between the sheet music and his fingers, ensuring he played each note correctly	She smiled at the audience, her saxophone gleaming under the stage lights as she played a upbeat tune	
They play a glissando by sliding their finger across the keys They keep their hands steady for a long, sustained note te tilts the trombone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	He slides his left hand along the strings to change the pitch	
They keep their hands steady for a long, sustained note He tilts the trombone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	They play a glissando by sliding their finger across the keys	
He tilts the trombone up for a high note The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	They keep their hands steady for a long, sustained note	
The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	He tilts the trombone up for a high note	
	The musician's eyes darted between the sheet music and his fingers, ensuring he played each note correctly	

Table 2. Example sentences from each OOD topic.

	Train				Test			
	Docitivo	In-Domain	Out-of-Domain	Docitivo	In-Domain	Out-of-Domain		
	rositive	Negative	Negative	Fositive	Negative	Negative		
QVHighlights [2]	7218	7218	7230	1550	1550	1550		
Charades-STA [1]	12404	12404	7230	3720	3720	1550		

Table 3. Numbers of positive and negative queries used for QVHighlights and Charades-STA.

#### **3.3. Implementation Details**

**UniVTG** For QVHighlights, we use loss weightings of  $\lambda^+ = 1$ ,  $\lambda_{ID}^- = 0.1$ ,  $\lambda_{OOD}^- = 0.1$ , and  $\lambda_p = 1$ , while for Charades-STA, we adjust  $\lambda_{ID}^- = 0.5$ ,  $\lambda_{OOD}^- = 0.5$ . The remaining loss weightings are retained from QVHighlights and Charades-STA training defaults in UniVTG. For negative queries the cosine similarity loss weighting  $\lambda_s^-$  is set equal to the intra video saliency loss weighting.

**QD-DETR** Loss weightings of  $\lambda^+ = 1$ ,  $\lambda_{ID}^- = 0.05$ ,  $\lambda_{OOD}^- = 0.05$ , and  $\lambda_p = 1$  are used for both QVHighlights and Charades-STA. For QVHighlights,  $\lambda_s^- = 1$  while for

Charades-STA,  $\lambda_s^- = 4$ .

**CG-DETR** The same weightings are used as in QD-DETR except  $\lambda_{ID}^- = 0.1$ ,  $\lambda_{OOD}^- = 0.1$  for both datasets. All other loss weightings retain their default values.

## 4. OOD Generalisability

To test the generalisability of the negative-aware approach for OOD query sentences, we test the UniVTG-NA model on OOD sentences from another scenario on which the model has not been trained. This scenario is 'musician performances' (see sample sentences in Table 2). The re-

	QVHighlights			Charades-STA				
			Rejection Acc. (%)				Rejectio	on Acc. (%)
Method	R1@0.5	R1@0.7	ID	OOD	R1@0.5	R1@0.7	ID	OOD
UniVTG [3] SVM	63.48 ( <b>-3.87</b> )	49.87 (-2.78)	94.77	97.74	53.47 (-6.75)	33.49 ( <b>-5.06</b> )	35.89	50.40
CG-DETR [4] SVM	62.84 ( <b>-4.26</b> )	50.26 (-3.29)	95.55	95.41	43.79 (-13.74)	28.44 (-7.23)	82.90	74.52
QD-DETR [5] SVM	48.26 (-13.74)	37.42 ( <del>-8.84</del> )	96.32	95.09	46.67 (-12.44)	30.05 ( <b>-6.70</b> )	76.91	81.59

Table 4. Results of training an SVM on top of the saliency score outputs of UniVTG, CG-DETR and QD-DETR.



Figure 1. The classification head for QD-DETR-NA and CG-DETR-NA takes as input a concatenation of saliency and indicator scores, which are then passed through a recurrent layer and a feed forward layer before producing a single value output for classification.

jection accuracy results are shown in Table 5. The rejection accuracy remains high for both datasets, demonstrating that the model is capable of generalising to other OOD scenarios.

	Rejection Acc. (%)				
Method	QVHighlights	Charades-STA			
UniVTG-NA	99.8	93.8			

Table 5. Rejection accuracy results for UniVTG-NA on the unseen OOD category of 'musician performance'.

# 5. UniMD

To further motivate the need for negative-aware training for the task of Negative-Aware Video Moment Retrieval, we investigate the output produced by UniMD [6], a recent SOTA method which only produces indicator scores with no saliency scores. We plot histograms of the output scores for positive and in-domain negative sentences for Charades-STA and ActivityNet-Captions, as in Figure 2. There is significant overlap between the positive and negative distributions which shows that the model is not designed to handle negative rejection. This further motivates the need for models which are specifically trained to carry out negative rejection alongside moment retrieval.

#### 6. Qualitative Results

We provide further qualitative results from UniVTG-NA on the QVHighlights and Charades-STA datasets in Figure 3 & 4. The model frequently successfully localises the positive sentences and rejects the negative sentences. Failure cases are included in the bottom right of each set of examples. The failure case in Figure 3 is a case of UniVTG-NA rejecting a positive sentence, while in Figure 4 the model fails to reject an ID negative sentence.

## References

- Jiyang Gao, Chen Sun, Zhenheng Yang, and Ram Nevatia. Tall: Temporal activity localization via language query. In *ICCV*, 2017. 1, 3
- [2] Jei Lei, Tamara L. Berg, and Bansal. Qvhighlights: Detecting moments and highlights in videos via natural language queries. In *NeurIPS*, 2021. 1, 3
- [3] Kevin Qinghong Lin, Pengchuan Zhang, Joya Chen, Shraman Pramanick, Difei Gao, Alex Jinpeng Wang, Rui Yan, and Mike Zheng Shou. Univtg: Towards unified video-language temporal grounding. In *ICCV*, 2023. 1, 2, 4
- [4] WonJun Moon, Sangeek Hyun, SuBeen Lee, and Jae-Pil Heo. Correlation-guided query-dependency calibration in video representation learning for temporal grounding. *CoRR abs/2311.08835*, 2024. 1, 4
- [5] WonJun Moon, Sangeek Hyun, SangUk Park, Dongchan Park, and Jae-Pil Heo. Query-dependent video representation for moment retrieval and highlight detection. In *CVPR*, 2023. 1, 4
- [6] Yingsen Zeng, Yujie Zhong, Chengjian Feng, and Lin Ma. Unimd: Towards unifying moment retrieval and temporal action detection. *ECCV*, 2024. 1, 4



Figure 2. Histograms of prediction (indicator) scores for positive and in-domain negative queries produced by the UniMD model.



Figure 3. Qualitative results from UniVTG-NA on QVHighlights.



Figure 4. Qualitative results from UniVTG-NA on Charades-STA.