Prompt ID	Chunk Size (frames)	F1@0.05	F1@0.1	F1@0.15	F1@0.2	F1@0.25	F1@0.3	F1@0.35	F1@0.4	F1@0.45	F1@0.5	Average
1	4	0.223	0.336	0.385	0.425	0.453	0.474	0.474	0.490	0.498	0.498	0.426
1	8	0.262	0.375	0.417	0.464	0.492	0.511	0.525	0.539	0.548	0.548	0.468
1	12	0.227	0.310	0.408	0.439	0.455	0.475	0.481	0.486	0.486	0.501	0.427
1	16	0.194	0.291	0.349	0.406	0.440	0.457	0.474	0.474	0.474	0.486	0.405
2	4	0.179	0.222	0.253	0.263	0.273	0.281	0.283	0.286	0.289	0.293	0.262
2	8	0.244	0.295	0.338	0.360	0.370	0.383	0.385	0.385	0.393	0.393	0.355
2	12	0.272	0.348	0.387	0.421	0.433	0.445	0.448	0.454	0.454	0.460	0.412
2	16	0.263	0.381	0.420	0.466	0.488	0.502	0.505	0.516	0.527	0.527	0.459
3	4	0.177	0.210	0.234	0.246	0.256	0.260	0.262	0.263	0.265	0.265	0.244
3	8	0.231	0.290	0.334	0.349	0.367	0.378	0.382	0.387	0.393	0.393	0.351
3	12	0.264	0.353	0.405	0.424	0.435	0.441	0.457	0.466	0.466	0.466	0.418
3	16	0.265	0.378	0.439	0.472	0.498	0.517	0.523	0.540	0.543	0.546	0.472
4	4	0.228	0.289	0.346	0.355	0.367	0.384	0.404	0.407	0.410	0.413	0.360
4	8	0.216	0.302	0.350	0.381	0.401	0.412	0.422	0.432	0.443	0.449	0.381
4	12	0.206	0.299	0.374	0.404	0.452	0.475	0.482	0.497	0.501	0.505	0.419
4	16	0.215	0.317	0.366	0.427	0.459	0.476	0.488	0.512	0.533	0.533	0.433
5	4	0.190	0.262	0.322	0.342	0.355	0.366	0.372	0.388	0.388	0.391	0.338
5	8	0.234	0.330	0.371	0.399	0.416	0.440	0.450	0.457	0.467	0.471	0.403
5	12	0.226	0.327	0.397	0.440	0.471	0.490	0.525	0.537	0.549	0.549	0.451
5	16	0.268	0.386	0.443	0.496	0.526	0.553	0.561	0.575	0.583	0.588	0.498

Table 1: **F1** metrics across different prompt configurations and chunk sizes

Prompt ID	Chunk Size (frames)	Precision@0.05	Precision@0.1	Precision@0.15	Precision@0.2	Precision@0.25	Precision@0.3	Precision@0.35	Precision@0.4	Precision@0.45	Precision@0.5	Average
1	4	0.189	0.285	0.326	0.361	0.385	0.402	0.402	0.416	0.423	0.423	0.361
1	8	0.250	0.357	0.397	0.442	0.469	0.487	0.500	0.513	0.522	0.522	0.446
1	12	0.239	0.326	0.429	0.462	0.478	0.500	0.505	0.511	0.511	0.527	0.449
1	16	0.231	0.347	0.415	0.483	0.524	0.544	0.565	0.565	0.565	0.578	0.482
2	4	0.108	0.134	0.152	0.158	0.164	0.169	0.170	0.172	0.174	0.176	0.158
2	8	0.164	0.198	0.227	0.242	0.249	0.257	0.259	0.259	0.264	0.264	0.238
2	12	0.197	0.251	0.279	0.303	0.312	0.321	0.323	0.328	0.328	0.332	0.297
2	16	0.206	0.298	0.329	0.365	0.382	0.393	0.396	0.404	0.412	0.412	0.360
3	4	0.104	0.123	0.137	0.144	0.150	0.152	0.153	0.154	0.155	0.155	0.143
3	8	0.149	0.187	0.215	0.225	0.236	0.243	0.246	0.249	0.253	0.253	0.226
3	12	0.184	0.245	0.281	0.294	0.302	0.306	0.317	0.323	0.323	0.323	0.290
3	16	0.197	0.281	0.327	0.351	0.370	0.385	0.389	0.401	0.404	0.406	0.351
4	4	0.161	0.204	0.245	0.251	0.259	0.271	0.286	0.288	0.290	0.292	0.255
4	8	0.166	0.232	0.268	0.292	0.308	0.316	0.324	0.332	0.339	0.345	0.292
4	12	0.166	0.241	0.301	0.325	0.364	0.383	0.389	0.401	0.404	0.407	0.338
4	16	0.183	0.270	0.311	0.363	0.391	0.405	0.415	0.436	0.453	0.453	0.368
5	4	0.132	0.182	0.224	0.237	0.247	0.254	0.258	0.270	0.270	0.272	0.234
5	8	0.179	0.253	0.285	0.306	0.319	0.338	0.346	0.351	0.359	0.361	0.310
5	12	0.186	0.270	0.328	0.363	0.389	0.405	0.434	0.444	0.453	0.453	0.373
5	16	0.241	0.348	0.399	0.447	0.474	0.498	0.506	0.518	0.526	0.530	0.449

Table 2: **Precision** metrics across different prompt configurations and chunk sizes

Prompt ID	Chunk Size (frames)	Recall@0.05	Recall@0.1	Recall@0.15	Recall@0.2	Recall@0.25	Recall@0.3	Recall@0.35	Recall@0.4	Recall@0.45	Recall@0.5	Average
1	4	0.271	0.409	0.468	0.517	0.552	0.576	0.576	0.596	0.606	0.606	0.518
1	8	0.276	0.394	0.438	0.488	0.517	0.537	0.552	0.567	0.576	0.576	0.492
1	12	0.217	0.296	0.389	0.419	0.433	0.453	0.458	0.463	0.463	0.478	0.407
1	16	0.167	0.251	0.300	0.350	0.379	0.394	0.409	0.409	0.409	0.419	0.349
2	4	0.527	0.655	0.744	0.773	0.803	0.828	0.833	0.842	0.852	0.862	0.772
2	8	0.478	0.576	0.660	0.704	0.724	0.749	0.754	0.754	0.768	0.768	0.694
2	12	0.443	0.567	0.631	0.685	0.704	0.724	0.729	0.739	0.739	0.749	0.671
2	16	0.365	0.527	0.581	0.645	0.675	0.695	0.700	0.714	0.729	0.729	0.636
3	4	0.606	0.719	0.803	0.842	0.877	0.892	0.897	0.901	0.906	0.906	0.835
3	8	0.517	0.650	0.749	0.783	0.823	0.847	0.857	0.867	0.882	0.882	0.786
3	12	0.473	0.631	0.724	0.759	0.778	0.788	0.818	0.833	0.833	0.833	0.747
3	16	0.404	0.576	0.670	0.719	0.759	0.788	0.798	0.823	0.828	0.833	0.720
4	4	0.389	0.493	0.591	0.606	0.626	0.655	0.690	0.695	0.700	0.704	0.615
4	8	0.310	0.433	0.502	0.547	0.576	0.591	0.606	0.621	0.635	0.645	0.547
4	12	0.271	0.394	0.493	0.532	0.596	0.626	0.635	0.655	0.660	0.665	0.553
4	16	0.261	0.384	0.443	0.517	0.557	0.576	0.591	0.621	0.645	0.645	0.524
5	4	0.340	0.468	0.576	0.611	0.635	0.655	0.665	0.695	0.695	0.700	0.604
5	8	0.335	0.473	0.532	0.571	0.596	0.631	0.645	0.655	0.670	0.675	0.578
5	12	0.286	0.414	0.502	0.557	0.596	0.621	0.665	0.680	0.695	0.695	0.571
5	16	0.300	0.433	0.498	0.557	0.591	0.621	0.631	0.645	0.655	0.660	0.559

Table 3: **Recall** metrics across different prompt configurations and chunk sizes

Note: Evaluation is performed on a subset of 100 videos.

Prompt Descriptions

Prompt 1:

- *Initial description:* Briefly describe the new event that just started.
- *Validity check:* Given the previous event description: {previous_description}, is this the same event or a new one? Respond ONLY with 'YES' if it's the same or 'NO' if it's a different event.
- *Update description:* Briefly describe the new event that just started.

Prompt 2:

- *Initial description:* Briefly describe the action/event that just started in the current video segment. Focus only on the main actions and the progress of the action, ignore meaningless detail to determine the current event. Be careful on their details. Describe ONLY what is happening in the video segment provided.
- *Validity check:* Your goal is to identify if the same action is ongoing in the current video segment compared to the previous one. Given the previous event description: {previous_description} Question: Is the current video segment still the same stage of the same event? As the last line answer ONLY with 'YES' or 'NO' with no extra characters.

- *Update description:* Briefly describe the action/event that just started in the current video segment. Focus only on the main actions and the progress of the action, ignore meaningless detail to determine the current event. Be careful on their details. Describe ONLY what is happening in the video segment provided.

Prompt 3:

- *Initial description:* Briefly describe the action/event that just started in the current video segment. Focus ONLY on the main character and the main actions. Describe only what's in the video segment
- *Validity check:* Your goal is to identify if the same action is ongoing in the current video segment compared to the previous one.

  Given the previous event description: {previous_description} Question: is this the same event or not? As the last line answer ONLY with 'YES' or 'NO'.
- *Update description:* Briefly describe the action/event that just started in the current video segment. Focus ONLY on the main character and the main actions. Describe only what's in the video segment

Prompt 4:

- *Initial description:* Provide a concise description of the main action in the current video segment. Focus only on the movements.
- *Validity check:* Given the previous event description: {previous_description}, is this the same event? Respond ONLY with 'YES' or 'NO'.
- *Update description:* Provide a concise description of the main action in the current video segment. Focus only on the movements.

Prompt 5:

- *Initial description:* Provide a concise description of the main action in the current video segment. Focus on the substeps of the action.
- *Validity check:* Given the previous event description: {previous_description}, is this the same event? Respond ONLY with 'YES' or 'NO'.
- *Update description:* Provide a concise description of the main action in the current video segment. Focus on the substeps of the action.