Reviews for Proposal 6

Reviewer 1

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
4: Top 25% (Excellent): The report covers existing work very well, demonstrating strong grasp and relevance, with minor areas for improvement.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
Innovation: This project excels in its clarity, ambition, and technical depth. The team shows a mature understanding of computer vision and deepfake detection challenges. The inclusion of both frame based and spatiotemporal models demonstrates awareness of tradeoffs between efficiency and accuracy. Relevance: The work aligns perfectly with current issues in AI-generated media and misinformation, tying directly to topics in machine learning, computer vision, and ethics. Methodology: The methods are well thought out and technically rigorous with its detailed training and calibration pipeline, evaluation plan (cross-generator robustness), and real world alignment with misinformation mitigation. Clarity: The writing is concise, logically structured, and easy to follow. The rationale behind each modeling decision is explained clearly, and the team provides a detailed four-week timeline with concrete milestones.
Weaknesses:
Vagueness: While the proposal is very strong, some aspects such as the calibration cost function and human validation threshold could benefit from more explicit mathematical or procedural detail. Methodological Issues: There's a potential risk of overfitting or model bias if synthetic data sources are not sufficiently diverse. Feasibility: Training heavy spatiotemporal models like SlowFast may exceed available compute or time constraints. Another possible challenge is balancing data diversity across generator types to avoid overfitting to specific synthesis families. Potential Risks: Ethical concerns around using real creator videos for "real" data should be addressed, even if the content is public. The team should also ensure all synthetic data sources comply with platform guidelines.
Ideas for improvement:
Clarification: Clarify how human-in-the-loop or abstention thresholds will be evaluated quantitatively. Addressing Limitations: Discuss data licensing/ethical considerations if scraping from platforms.
Additional Comments:
nan
Grade Prediction: A
Estimate of final project vs proposal: 4 – High Alignment: The project generally follows the proposal with only minor changes or refinements. The core goals and methods remain consistent.

Reviewer 2

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing

knowledge.
Does the report sufficiently cover existing work?
5: Top 10% (Exceptional): The report provides a comprehensive and critical review of existing work, showcasing exceptional understanding and analysis.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
The amount of detail of this project is really impressive. I also think trying to tackle a real world problem that everyone is facing is great. This proposal shows a great understanding of the technical depth and awareness with the details.
Weaknesses:
I think a roadblock could be the complexity of proposal due to the limited of time that we have until the end of the semester. Also, a possibility that could be a roadblock is the dataset that will be used. In the proposal it is mentioned scraping will be used but some websites, scraping violates it user agreement and could be banned.
Ideas for improvement:
I think explaining more on how you will obtain the dataset, could be beneficial for the next proposal. When the data is obtained adding a section for evaluation if there is any bias from you results. While all the techniques are great but if you narrow down to one approach it could be beneficial due to limited time.
Additional Comments:
The proposal is great and look forward to the results
Grade Prediction: A
Estimate of final project vs proposal: 5 – Very High Alignment: The project plan and execution appear almost identical to the original proposal. Clear evidence that the group will deliver what was proposed with minimal deviation.

Reviewer 3

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
4: Top 25% (Excellent): The report covers existing work very well, demonstrating strong grasp and relevance, with minor areas for improvement.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
The team’s approach is technically rigorous and well thought out, combining spatial and temporal modeling. The inclusion of advanced regularization techniques such as mixup, cutmix, and stochastic depth demonstrates a solid understanding of modern deep learning practices. The integration of post-hoc calibration and human-in-the-loop decision routing is exceptional.
Weaknesses:
Acquiring and preprocessing diverse short-form video datasets, especially synthetic data from multiple generation sources, could be time-consuming and raise issues related to quality consistency or licensing.
Ideas for improvement:
Exploring lightweight multimodal extensions that integrate audio signals could improve robustness to

cross-domain variations. Additionally, implementing qualitative visualization tools like saliency maps or attention heatmaps could help interpret the model's decision process and provide greater transparency.
Additional Comments:
nan
Grade Prediction: A
Estimate of final project vs proposal: 5 – Very High Alignment: The project plan and execution appear almost identical to the original proposal. Clear evidence that the group will deliver what was proposed with minimal deviation.

Reviewer 4

Is the problem well-specified, interesting, and ambitious?
4: Top 25% (Excellent): The problem is very well-specified, captivating, and shows a high level of ambition. It is interesting, relevant, and presents a significant challenge.
Does the report sufficiently cover existing work?
1: (Not Adequate): The report does not sufficiently cover existing work. For example, a quick Google/Google Scholar search shows that reasonably similar things have been attempted before, but they are not mentioned; or it is not clear how the proposed work meaningfully differs from existing work.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
Authenticity Understanding for Real vs Artificial Shortform Videos is innovative and much required in today's AI content-driven world. It is under-explored in research and very relevant to modern challenges like social media moderation and AI ethics. The research proposal has thoroughly lined out methodologies from baseline models to CNNs and transformers exploring different kinds of methodologies suitable to the problem. Evaluation does not just include functional but non functional requirements too like robustness, calibration and generalization. Proposal is well structured, clear and easy to follow with a well defined workflow.
Weaknesses:
Labels for the dataset are unclear in terms of the reliability of a video being real vs synthetic. The proposal does not clearly discuss the existing work and how it is different from the methods that have already been implemented. The models discussed in the proposal may be computationally heavy requiring more than expected time to train. Financial and logistical feasibility of gathering data synthetically and by scraping is a concern.
Ideas for improvement:
There should be a process specified to identify how a video will be classified and labelled as synthetic or real. There may be ethical and privacy concerns in the videos which could be addressed. Additionally, cross platform videos could be used to train the model to make it more robust and generalizable. Other than video, audio could also be integrated in the model as it can provide more information about the authenticity of the video. Future work could include doing a multimodal detection like audio, video, text etc.
Additional Comments:
The problem is a well defined one with implications in the real world. The proposal describes a well thought out and technically sound plan. The project can meaningfully contribute to fields like content moderation, AI ethics and media forensics. With some thought about data labelling and computational feasibility, this is a strong and meaningful project.
Grade Prediction: A
Estimate of final project vs proposal: 4 – High Alignment: The project generally follows the proposal with only minor changes or refinements. The core goals and methods remain consistent.

Reviewer 5

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
4: Top 25% (Excellent): The report covers existing work very well, demonstrating strong grasp and relevance, with minor areas for improvement.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
Clear product goal: calibrated detection to triage moderation. Strong baseline→advanced model ladder. Thorough training and robustness tactics. Explicit calibration and abstention policy. Cross-generator and cross-dataset evaluation. Concrete compute stack and weekly timeline.
Weaknesses:
Data acquisition and rights. The plan mentions scraping YouTube/TikTok. That can conflict with ToS and watermark policies. Potential demographic and content bias. Physiological cues can be brittle across lighting and demographics. Scope risk. Two model families plus calibration and robustness might stretch time if GPU access slips.
Ideas for improvement:
Lock one primary deliverable (e.g., the frame→temporal model) as the must-ship model. Treat the spatiotemporal model as stretch. Define a concrete success bar now (e.g., AUROC ≥ X on held-out generator; ECE ≤ Y after calibration; robust AUC drop ≤ Z after platform transcodes). Add a fairness/shift check. Report performance by content type and skin-tone/lighting proxies if available. Note limits if labels are absent.
Additional Comments:
nan
Grade Prediction: A to A-
Estimate of final project vs proposal: 5 – Very High Alignment: The project plan and execution appear almost identical to the original proposal. Clear evidence that the group will deliver what was proposed with minimal deviation.

Reviewer 6

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
4: Top 25% (Excellent): The report covers existing work very well, demonstrating strong grasp and relevance, with minor areas for improvement.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
I like how thorough the plan is

Weaknesses:
It just seems like there's a ton of things going on so I could see it be overwhelming at a point
Ideas for improvement:
I don't know much about generators or LLMS but I think that was good. Maybe they could mention more about their presentation.
Additional Comments:
Project looks really well done
Grade Prediction: A
Estimate of final project vs proposal: 3 – Moderate Alignment: The project retains the main ideas of the proposal but shows notable changes in scope, approach, or expected outcomes.

Reviewer 7

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
4: Top 25% (Excellent): The report covers existing work very well, demonstrating strong grasp and relevance, with minor areas for improvement.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
I believe the project's greatest strength lies first in its highly significant topic selection. With deepfakes and AI-generated videos rapidly proliferating across social platforms, the risks they pose for spreading misinformation have become a pressing reality. Therefore, the team's proposed “Short Video Authenticity Detection System” holds direct industry application value. Additionally, our course SI 670, titled Applied Machine Learning, aligns perfectly with its focus on solving real-world AI challenges, showcasing remarkable potential for societal impact. Second, the project plan demonstrates creativity. The team not only employs a frame feature fusion strategy combining ViT/EfficientNet with GRU/Transformer but also explores advanced video networks like SlowFast or X3D that explicitly model spatio-temporal information. Having not researched these models myself, this indicates the team has diligently studied various model-related knowledge outside of class and excels at integrating their strengths. Regarding feature dimension selection, the team focused on common physical/physiological flaws in generated videos, such as inconsistent microexpressions, abnormal blink frequencies, and motion trajectories violating physiological norms. This reflects the team's independent analysis, observation, research, and evaluation of such issues. I've only heard about physiological flaws in AI videos, like extra fingers. Movement trajectories defying common sense. However, the team also addressed micro-expression consistency and blink frequency. This not only enhances model interpretability but also aligns with existing research analyzing deepfake's inherent flaws, constituting clear and substantial supporting work. Most importantly, it demonstrates the team's determination to tackle complex models and intricate features—just addressing expression consistency alone makes me imagine the massive data volume required. Furthermore, the team's engineering implementation is exceptionally robust. They clearly define GPU resource requirements, employ mixed-precision training strategies, handle multi-source data format alignment, plan storage overhead, and establish a comprehensive schedule with clear and reliable task division. These elements demonstrate professional ML engineering thinking. While classroom instruction and textbooks often focus solely on data preprocessing, model training, and validation sets, this project's consideration of GPU resources, storage capacity, and storage overhead planning clearly indicates that the team includes individuals with practical experience applying ML in industrial settings. This resonates deeply with me, as GPU and

storage requirements typically involve rough estimates based on model and parameter specifications. The project's strong feasibility sets it apart—it transcends an academic concept to become a genuinely deployable system. Finally, the proposal maintains a clear structure with rigorous logic and well-defined priorities. It effectively communicates the problem motivation, methodology, and evaluation approach, demonstrating excellent communication skills.

Weaknesses:

However, given the immense complexity of this project, I first questioned their feasibility. Upon investigation, I learned that models like SlowFast/X3D are extremely GPU-intensive, with video processing I/O becoming the bottleneck. This often necessitates cloud computing as an auxiliary solution. Even with GreatLake, I remain concerned about the team's budget and time constraints. Moreover, the recent release of SORA seems to signal a new era for video generation models. While the existing related work cites SlowFast, X3D, and rPPG, it lacks a comprehensive review of recent adversarial studies on generative video models. It also fails to adequately discuss key trends in fake detection development from 2024–2025. I merely suggest this could be included in the related work section, though its actual implementation would be extremely challenging. Secondly, regarding feature selection. While micro-expression analysis and blinking are excellent feature points, I believe both data collection and quantification are overly complex. This often requires frame-by-frame data logging. Technically, rPPG and blink detection demand reliable tracking and are susceptible to lighting and occlusion issues. Microexpression distortion detection relies on facial cropping and keypoint extraction. These feature extraction methods, when using blinking, microexpressions, or rPPG, depend on tracking success rates. Tracking failures can significantly degrade overall results. Additionally, there's the fairness issue discussed in class. YouTube/TikTok data may involve user portrait rights, and platform data collection must comply with regulations. Misjudgments by algorithms can damage real users' reputations. Binary classification of “real vs. fake” inherently involves credibility judgments. Improper decision thresholds may cause real videos to be flagged as fake, harming genuine creators. Overall, I believe the primary concern lies in the level of completion. Given that the research tackles a cutting-edge, highly challenging problem in the industry, I deeply admire the team's courage. However, I also have some concerns about the level of completion

Ideas for improvement:

This is purely my personal opinion!!!!! I still have some doubts about the project's feasibility and implementation potential, so everything I say below is based on that perspective. If this project can be implemented and completed exceptionally well, then I don't believe I have the standing to question their qualifications here. First, to enhance the project's clarity and feasibility, the team should consider narrowing the initial scope to ensure core functionality is delivered. For instance, starting with a frame-to-sequence model pipeline and introducing the full spatio-temporal model only after establishing a stable baseline would help manage GPU resource and time constraints. To enhance robustness, feature design could incorporate fallback strategies for facial tracking failures—such as leveraging global motion anomaly cues or audiovisual synchronization analysis—ensuring the system functions reliably even with low-quality or occluded videos common on short-video platforms. Finally, I strongly recommend establishing a prioritized delivery roadmap—progressing from a minimum viable system to robustness enhancements, then to physiological feature analysis—to balance technical ambition with reliability. Any remaining advanced components (e.g., rPPG analysis) can serve as future research directions beyond core submission milestones. Achieving a minimum viable system ensures at least partial project completion. If this project is implemented, I recommend expanding the relevant work section to cover recent advancements in video diffusion techniques and commercial AI video generation tools (e.g., Sora, Runway Gen-2, Luma Dream Machine). This would directly address cutting-edge industrial challenges and tackle the latest issues.

Additional Comments:

Overall, I find this project both highly relevant and profoundly insightful. It addresses a significant and pressing real-world challenge that aligns closely with applied machine learning practices. The proposal demonstrates clear strengths in creativity, technical ambition, and engineering detail. While feasibility risks exist and synthetic video generation technology evolves rapidly, the structured plan gives me confidence that the team will deliver valuable outcomes. I sincerely commend the team's vision and

courage in selecting such a complex and far-reaching research domain. As a practitioner equally focused on trustworthy AI and combating misinformation, this project aligns strongly with my personal values. I look forward to seeing how this model performs on actual short-form video content. My sole concern is: If GPU resources prove tighter than anticipated, what is the minimum deliverable milestone? In summary, this stands as one of the most technically impressive and socially significant proposals I've reviewed. It also ranks among the most hardcore classroom projects I've encountered.
Grade Prediction: A
Estimate of final project vs proposal: 2 – Low Alignment: The project diverges significantly from the proposal in several key areas (e.g., focus, methods, deliverables). Some connection remains, but it’s weak.

Reviewer 8

Is the problem well-specified, interesting, and ambitious?
5: Top 10% in class (Exceptional): The problem is exceptionally well-specified, highly interesting, and extremely ambitious. It demonstrates innovation, deep understanding, and a drive to challenge existing knowledge.
Does the report sufficiently cover existing work?
5: Top 10% (Exceptional): The report provides a comprehensive and critical review of existing work, showcasing exceptional understanding and analysis.
Are Methods, Datasets, and Evaluation well-specified and appropriate?
5: Top 10% (Exceptional): The methods, datasets, and evaluation are exceptionally well-specified and highly appropriate, demonstrating expertise and meticulous planning.
Strengths:
I think the topic is very innovative and relevant especially given how rampant and unchecked AI videos are circulating these days. It's highly necessary for regulatory technologies to be ahead of generative tech instead of failing to catch up and leading to ethical and legal issues.
Weaknesses:
I'm not familiar with video prompt engineering skills of the authors of this proposal so this might not be applicable: synthetic videos generated for the purpose of this project might not be as realistic or of the same style as hyper-realistic synthetic videos generated by content creators on social media, which could affect effectiveness of the model.
Ideas for improvement:
I wonder if it'll be helpful to focus on a specific type of video (whether it's style, topic, or by another criteria) for the purpose of this project to improve model accuracy?
Additional Comments:
nan
Grade Prediction: 95
Estimate of final project vs proposal: 4 – High Alignment: The project generally follows the proposal with only minor changes or refinements. The core goals and methods remain consistent.