A machine learning model to automatically identify fast-paced online videos for children

May 2025

Claire Christensen, Ph.D., Anirban Roy, Ph.D.,
Madeline Cincebeaux, BS , Ramneet Kaur, Ph.D., SRI Education
Jenny Radesky, M.D., Tiffany Munzer, M.D., University of Michigan Medical Center

Exposure to fast-paced videos may be harmful for children

Most young children watch videos online (Mann et al., 2025), and many of these videos are **fast-paced**. In a dataset of 426 online videos watched by 47 children from ages 0-35.9 months, researchers found that 73% of videos were fast-paced, characterized by rapid cuts, frequent camera changes, or novel visual stimuli (Henderson et al., 2024).

Exposure to fast-paced videos has been linked to **adverse outcomes** among children aged 8 and under, including diminished executive function and attention (e.g., Kostyrka-Allchorne et al., 2017; Lillard & Peterson, 2011; Cooper et al., 2009; Rose et al., 2022). Fast paced videos are theorized to exert "bottom-up" attentional control for children and supplant opportunities for children to develop "top-down" executive functioning skills that often occur with child-led, autonomy supportive activities such as free play (Lillard et al., 2015). For example, preschoolers who watched 9 minutes of a fast-paced cartoon performed significantly worse on executive function tasks than those who engaged in drawing activities, and marginally worse than those who watched educational television (Lillard and Peterson, 2011). Another study found that 3-year-olds showed more attention and effort on a subsequent problem-solving task after watching slower-paced content, as compared to fast-paced content (Rose et al., 2022). And observational studies have linked early exposure to fast-paced content with later attentional problems (Zimmerman & Christakis, 2007; Christakis et al., 2004).

Suggested Citation: Christensen, C., Roy, A., Cincebeaux, M., Kaur, R., Radesky, J., & Munzer, T. (2025). *A machine learning model to automatically identify fast-paced online videos for children.* Menlo Park, CA: SRI Education.

While executive function is the most studied outcome of exposure to fast-paced videos, researchers have also hypothesized that exposure to fast-paced videos may contribute to other negative outcomes, including **problematic media use** (PMU). Fast pacing is one of several features categorized as persuasive design, which also includes bright colors, loud sound effects, intermittent rewards, wish-fulfillment themes, and intense emotions that capture and sustain child attention (Coyne et al., 2024; Domoff et al., 2020; Radesky, 2018; Sakamoto et al., 2022). According to the Interactional Theory of Childhood Problematic Media Use (Domoff et al., 2020), these features may promote PMU by increasing sustained engagement and reducing natural opportunities for natural disengagement. Fast pacing, in particular, may encourage longer duration of game play or video viewing because children look longer at content that includes cuts and movement (Schmitt et al., 1999).

We need more research on fast-paced content in online videos

Existing research on fast-paced video exposure does not reflect young children's current media habits. Children ages birth–8 now spend about 60% of their daily screen time watching online videos, such as those on YouTube (Mann et al., 2025). **Yet no studies to date have examined the effects of exposure to fast-paced content in online videos.** Instead, prior work largely relied on commercially available television shows (e.g., SpongeBob SquarePants; Lillard et al., 2015) or experimentally manipulated video clips (e.g., Kostyrka-Allchorne et al., 2017). These types of media may differ from online videos in important ways. For instance, online videos may be less likely to use narrative structures and characters to support educational content (Christensen & Cincebeaux, 2024); may be shorter in duration than typical TV shows; and, given that many online videos are user-generated, may include more scene changes or camera movement. Studies of YouTube algorithms suggest that creators use attention-grabbing design features to engage young viewers with their content (Radesky et al., 2024). Given these differences, it is critical to examine how the pacing of online videos, specifically, affects young children. This research is essential to provide guidance for parents, educators, and healthcare providers about whether to limit children's exposure to fast-paced online video content.

There is no single, straightforward definition for fast-paced content

Studies use a wide range of metrics to define fast-paced content. They include average shot length, number of edits per minute, rate of new content introduction, frequency of scene or character changes, and the speed and intensity of visual and auditory stimuli (see Table 1). Some studies do not define fast pace at all.

Motion: A new approach to identifying fast-paced videos

A clear, scalable method for identifying fast-paced videos is necessary to accurately assess whether exposure to fast-paced online videos is associated with negative outcomes in children. Manual human annotation, the current standard practice, is not a viable solution for today's media landscape. The sheer volume and diversity of online video content make manual coding impractical, and achieving reliability among human annotators when rating video pace is notoriously difficult, in part because pacing can vary widely throughout a video.

Table 1. Research studies examining fast-paced video content for children

Citation	Participant Age Range	Study Design	Definition of Fast Pace	Video Type	Findings
Anderson et al., 1977	4-year-olds	Random assignment to fast-paced Sesame Street, slow-paced Sesame Street, or story-reading	Frequency of edits, scene changes, motion, and audio changes; use of lively audio	40-min Sesame Street episode	Controlling for content, pace alone was not linked to changes in impulsivity, attention, or play behaviors.
Cooper et al., 2009	4–7 years	Random assignment to fast- or slow-paced video	Fast-edit: Scene changes every 4 seconds on avg; Slow-edit: scene changes every 15 seconds on avg	3.5-min edited video	4-year-olds in the slow-edit group showed better orienting; the effect reversed at age 6. In all age groups, children who watched a slow-edited film made more errors on an attention network task.
Gola & Calvert, 2011	6, 9, and 12 months	Randomized to fast- or slow-paced video sets	Number of scene/character changes per minute; fast-paced had higher rate	5-min segments from commercial infant DVDs	Attention was higher during fast-paced DVDs, especially for 6- and 9-month-olds.
Goodrich et al., 2009	0–72 months	Descriptive analysis of media formal features	Scene and character/object changes per minute	59 commercial infant/toddler DVDs	Most programs used perceptually salient features like fast pacing.
Henderson et al., 2024	0–35.9 months	Content analysis of videos from the 2020 Common Sense Census	Fast cuts or new concepts introduced faster than once every 20 sec	426 videos watched by 47 children	73% of videos were fast-paced.
Lillard & Peterson, 2011	4-year-olds	Random assignment to fast-paced cartoon, educational cartoon, or drawing	Fast-paced: scene changes every 11 sec; educational: scene changes every 34 sec	Fast-paced: SpongeBob; slow-paced: Caillou	Fast-paced group performed worse on executive function tasks, controlling for age, attention, and TV exposure.
Rose et al., 2022	3- and 4-year-olds	Children visited twice at home; shown an episode that was faster or slower paced	Higher frequency of cuts, scene changes, camera angles, number of characters per episode, and sound effects	Two 15-minute episodes of the TV program Postman Pat, one slow-paced and one fast-paced.	Children attended more to the fast-paced program than to the slow-paced one. There was a significant age-by-video-pace interaction for attention, such that 3-year-olds exhibited greater attention on a task after watching the slow-paced video, but there were no differences for 4-year-olds.

To address these challenges, we have developed an approach to measure video pace via optical flow—the amount of motion in video pixels across frames of a video. This metric captures the speed and direction at which objects move in a video. It is sensitive to the dynamic qualities that many prior definitions of fast-paced media have emphasized. By **focusing directly on motion** rather than on proxies like shot length, our measure provides a more consistent and scalable approach to identifying fast-paced content across the broad range of videos children watch today.

Building a machine learning tool to automatically detect fast-paced videos for young children

Training the model on expert-annotated videos watched by real kids

We trained and evaluated our machine learning model using a dataset of YouTube videos that had been annotated by trained and reliable human coders. These annotations were drawn from a previously published study of media content viewed by infants and toddlers (Henderson et al., 2024). Participants were 47 children between 0 and 35.9 months of age whose caregivers participated in the 2020 Common Sense Census. These caregivers submitted their child's most recent YouTube viewing history, resulting in an initial pool of 470 videos. After removing duplicates, unavailable links, and non-English content, 426 unique videos were included in the final coded sample. These videos were systematically annotated for features such as educational content, pacing, and comprehension-aiding techniques using a coding scheme adapted from prior work on infant-directed media. Due to subsequent video availability issues, we were able to download and process 415 of the 426 videos for use in our machine learning pipeline. The annotated dataset thus reflects a diverse and ecologically valid sample of content actually watched by infants and toddlers on YouTube.

Research question

We used these data to answer this question:

> To what extent can a machine learning model accurately detect video pace, as compared to trained human annotators?

Training a machine learning model to detect fast-paced videos using motion

We used the annotated Henderson (2024) dataset to train and test a machine learning model to detect fast-paced video content. We first measured the motion in a video using optical flow and then used a machine learning model to determine whether a given video is fast- or slow-paced based on motion.

To quantify the motion in a video, we used a technique from computer vision called **optical flow**. This method captures the movement of each pixel between consecutive frames in a video. Essentially, it tracks how much and in what direction each part of the image shifts over time. The average motion values between the frames in a video is represented by a 1-D time series. To compute metrics that accurately represent these data, we used a sliding window technique to divide the time series into small overlapping segments. For each segment, we calculated simple features (like the average amount of motion), which summarizes local patterns in the video's movement. This approach is helpful because it preserves information about how motion changes over time, rather than collapsing the video into a single summary number.

We then trained a **1-D convolutional neural network** (CNN) model, a type of model well-suited for analyzing time series data, to understand how these motion metrics are related to human-perceived pace.

Human annotations of pacing (see above) served as the "ground truth" for our machine learning model, meaning that they taught the model what fast and slow videos look like in terms of their motion patterns. The model learns to recognize patterns in the motion signals in the video that are predictive of fast or slow pacing as annotated by humans. Once trained, the model can classify new videos as fast- or slow-paced based on their motion patterns alone, without the need for human annotation.

We follow an iterative process to train and test the machine learning model. Each iteration involves the following process:

1. **Split:** We randomly assign 75% of the videos to a model training set and 25% to a model testing set. This allows us to test the model's performance on videos that it did not use for training.
2. **Train:** We train the model on the 75% training set.
3. **Test:** We ask the model to independently classify the 25% testing set.
4. **Compare:** We compare the model's video classifications with the human annotator classifications.
5. **Repeat:** To produce more reliable results, we train and test the model on three different 75%/25% training/testing splits by repeating Steps 1–4 three times.
6. **Average:** We then average the results of the three testing rounds to produce one overall precision score for video pace.

Evaluating the model using precision and accuracy by chance

When we evaluate the model's video classifications of whether a video is fast- or slow-paced, we use two comparisons. First, we compare the model's annotations to human annotations. The metric we use for this is **precision**. Precision tells us how many of the videos our model labeled as having fast-paced content matched the human annotations. For example, if our model said that 10 videos were fast-paced, but human annotations indicated that only 7 were fast-paced, the precision would be 7 out of 10, or 70%. We focus on precision because we want to be confident that the videos the model identifies as fast-paced actually are.

To better understand how well our model performs, we also compare its precision to a benchmark called **accuracy by chance**. This helps us judge that the model's evaluation is smarter than random guessing. Accuracy by chance reflects the base rate of fast-paced videos in our dataset. For example, if 60% of the videos are fast-paced, then randomly labeling videos as fast or slow would be right about 60% of the time, just by chance. That's our baseline. To estimate this, we simulate what would happen if we randomly guessed whether each video was fast or slow. We do this many times and take the average accuracy. If the model's precision is much higher than this baseline, the model is finding meaningful patterns beyond chance.

The model was 85% accurate for detecting fast-paced videos

To evaluate the model's precision, we compared its predictions to those of trained human annotators. The final model correctly identified slow-paced videos 83% of the time and fast-paced videos 85% of the time.

For example, if the model labels 100 videos as fast-paced, human raters agree on 85 of those classifications. These accuracy levels significantly exceed chance performance (Table 2), indicating that the model captures meaningful motion patterns associated with pace.

The slightly higher accuracy for fast-paced content may be due to the greater number of fast-paced videos in the training set (221 fast-paced vs. 85 slow-paced). This imbalance could have helped the model better learn the distinguishing features of fast-paced videos. Nevertheless, the model demonstrated strong performance for both classes and can be used to automatically detect video pacing in large datasets without requiring human annotation for each new video.

Table 2. Model precision and accuracy by chance for detecting video pace

Video pace	Training set	Model accuracy	Accuracy by chance
Slow	85	83%	27%
Fast	221	85%	72%

Laying the Groundwork for Scalable Media Content Research

This project demonstrates the feasibility of using machine learning to detect persuasive design features in children's media beginning with video pacing, one of the most foundational. We developed a machine learning model that analyzes motion cues in videos to classify content as fast- or slow-paced. Trained on human-annotated YouTube videos viewed by infants and toddlers, the model achieved 85% accuracy for fast-paced videos and 83% for slow-paced videos, which are significantly above chance and aligned with expert judgment.

Unlike traditional methods that measure pacing by counting scene changes, our approach focuses on movement onscreen and provides a complementary view of how fast-paced content engages young viewers. Although our model does not yet detect scene cuts, it captures the intensity and continuity of motion, which may be equally or more relevant to how children experience video content (Schmitt et al., 1999).

Our goal is to expand this work to detect additional persuasive design features, such as wish fulfillment themes, bright colors, loud sound effects, and intense emotion, and link these features to developmental outcomes. Our longer-term goal is to integrate these models into a longitudinal study tracking children's naturalistic media use and its impact on attention, executive function, and PMU.

This work is urgently needed. PMU is prevalent among young children (Swit et al., 2023), and persuasive design is implicated in its development (Domoff et al., 2020). To support children's well-being and inform evidence-based design, we must understand how specific features shape developmental outcomes. This project proves that it is possible to detect those features at scale. Our goal is to build on this progress to advance science, inform design, and support healthier digital experiences for young children.

References

Anderson, D. R., Levin, S. R., Lorch, E. P. (1977). The effects of TV program pacing on the behavior of preschool children. *AV Communication Review, 25*(3), 159–166. https://doi.org/10.1007/bf02769779

Christakis, D. A., Zimmerman, F. J., DiGiuseppe, D. L., & McCarty, C. A. (2004). Early television exposure and subsequent attentional problems in children. *Pediatrics, 113*(4), 708–713. https://doi.org/10.1542/peds.113.4.708

Christensen, C., & Cincebeaux, M. (2024). *YouTube's ABCs and 123s: Describing the quality of early literacy and math videos on YouTube.* Menlo Park, CA: SRI International.

Cooper, N. R., Uller, C., Pettifer, J., & Stolc, F. C. (2009). Conditioning attentional skills: Examining the effects of the pace of television editing on children's attention. *Acta Paediatrica, 98*(10), 1651–1655. https://doi.org/10.1111/j.1651-2227.2009.01377.x

Coyne, S. M., Holmgren, H. G., Shawcroft, J. E., Barr, R., Davis, E., Ashby, S., Stockdale, L., & Domoff, S. (2022). ABCs or attack–boom–crash? A longitudinal analysis of associations between media content and the development of problematic media use in early childhood. *Technology, Mind, and Behavior, 3*(4). https://doi.org/10.1037/tmb0000093

Domoff, S. E., Borgen, A. L., Radesky, J. S. (2020). Interactional theory of childhood problematic media use. *Human Behavior and Emerging Technologies, 2*(4), 343–353. https://doi.org/10.1002/hbe2.217

Gola, A. A. H., & Calvert, S. L. (2011). Infants' visual attention to baby DVDs as a function of program pacing. *Infancy, 16*(3), 295–305. https://doi.org/10.1111/j.1532-7078.2010.00051

Goodrich, S. A., Pempek, T. A., & Calvert, S. L. (2009). Formal production features of infant and toddler DVDs. *Archives of Pediatrics & Adolescent Medicine, 163*(12), 1151–1156. https://doi.org/10.1001/archpediatrics.2009.201

Henderson, D., Bailes, T., Sturza, J., Robb, M. B., Radesky, J. S., & Munzer, T. G. (2024). YouTube for young children: What are infants and toddlers watching on the most popular video-sharing app? *Frontiers in Developmental Psychology, 2.* https://doi.org/10.3389/fdpys.2024.1335922

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). The relationship between television exposure and children's cognition and behaviour: A systematic review. *Developmental Review, 44*(June 2017), 19-58. http://dx.doi.org/10.1016/j.dr.2016.12.002

Lillard, A. S., & Peterson, J. (2011). The immediate impact of different types of television on young children's executive function. *Pediatrics, 128*(4), 644–649. https://doi.org/10.1542/peds.2010-1919

Lillard, A. S., Drell, M. B., Richey, E. M., Boguszewski, K., & Smith, E. D. (2015). Further examination of the immediate impact of television on children's executive function. *Developmental Psychology, 51*(6), 792–805. https://doi.org/10.1037/a0039097

Mann, S., Calvin, A., Lenhart, A., and Robb, M.B. (2025). *The Common Sense census: Media use by kids zero to eight, 2025*. San Francisco, CA: Common Sense Media.

Radesky, J. S. (2018, February 6). *Persuasive Digital Design: Appealing to Adults, Problematic for Kids?* https://www.michiganmedicine.org/health-lab/persuasive-digital-design-appealing-adults-problematic-kids

Radesky, J., Bridgewater, E., Black, S., O'Neil, A., Sun, Y., Schaller, A., ... & Campbell, S. W. (2024). Algorithmic Content Recommendations on a Video-Sharing Platform Used by Children. *JAMA Network Open, 7*(5), e2413855-e2413855.

Rose, S. E., Lamont, A. M., Reyland, N. (2022). Watching television in a home environment: effects on children's attention, problem solving and comprehension. *Media Psychology, 25*(2), 208–233. https://doi.org/10.1080/15213269.2021.1901744

Sakamoto, N., Kabaya, K., & Nakayama, M. (2022). Sleep problems, sleep duration, and use of digital devices among primary school students in Japan. *BMC Public Health, 22,* 1006. https://doi.org/10.1186/s12889-022-13389-1

Schmitt, K. L., Anderson, D. R., & Collins, P. A. (1999). Form and content: Looking at visual features of television. *Developmental psychology, 35*(4), 1156.

Swit, C. S., Coyne, S. M., Shawcroft, J., Gath, M., Barr, R., Holmgren, H. G., & Stockdale, L. (2023). Problematic media use in early childhood: The role of parent-child relationships and parental wellbeing in families in New Zealand and the United States. *Journal of Children and Media, 17*(4), 443–466. https://doi.org/10.1080/17482798.2023.2230321

Zimmerman, F. J., & Christakis, D. A. (2007). Associations between content types of early media exposure and subsequent attentional problems. *Pediatrics, 120*(5), 986–992. https://doi.org/10.1542/peds.2006-3322

SRI Education, a division of SRI, is helping federal and state agencies, school districts, major foundations, nonprofit organizations, and international and commercial clients tackle some of the most complex issues in education to help students succeed. Our mission is **to reduce barriers and optimize outcomes for all children, youth, and families**. We do this by conducting high-quality research, supporting use of data and evidence, helping to strengthen state and local systems, and developing tools that improve teaching and accelerate and deepen learning. Our work covers a range of topics, including early learning and development, student behavior and well-being, teaching quality, digital learning, STEM and computer science, literacy and language arts, and college and career pathways.

SRI is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society.

Silicon Valley
(SRI Headquarters)
333 Ravenswood Avenue
Menlo Park, CA 94025
+1.650.859.2000
education@sri.com

Washington, DC
1100 Wilson Boulevard
Suite 2700
Arlington, VA 22209
+1.703.524.2053
www.sri.com/education-learning/

STAY CONNECTED