

Perancangan Sistem Klasifikasi Video Real-Time Berbasis Adaptation Layer dan State-Space Model pada Arsitektur LRCN = Design of a Real-Time Video Classification System Using an Adaptation Layer and State-Space Model in LRCN Architecture

Ahmad Rifqi Fadhlurrahman, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920571632&lokasi=lokal>

Abstrak

Dalam era digital, video menjadi bentuk konten utama di berbagai platform media sosial seperti TikTok. Dengan jutaan video yang diunggah setiap hari, diperlukan sistem klasifikasi otomatis yang mampu memproses dan mengkategorikan konten secara cepat dan efisien untuk mendukung moderasi konten. Penelitian ini bertujuan untuk merancang dan mengevaluasi sistem inferensi real-time berbasis deep learning untuk klasifikasi video TikTok, dengan fokus pada efisiensi sistem, akurasi model, serta kestabilan proses pelatihan. Metodologi yang digunakan melibatkan modifikasi arsitektur Long-term Recurrent Convolutional Network (LRCN) dengan menggantikan LSTM menggunakan State Space Model (SSM) sebagai drop-in replacement untuk pemodelan sekuens temporal. Arsitektur sistem terdiri dari CNN untuk ekstraksi fitur spasial, jaringan saraf 3-lapisan untuk koneksi antar modul, serta penggunaan ZeroMQ dan Docker untuk mengurutkan data dan isolasi proses inferensi. Evaluasi dilakukan pada skenario batch dan real-time, mencakup pengukuran akurasi, F1-score, jumlah parameter, latensi inferensi, dan kestabilan pelatihan. Hasil penelitian menunjukkan bahwa SSM dan LSTM memiliki akurasi klasifikasi yang setara (sekitar 76%), namun SSM memiliki parameter yang lebih sedikit (50% parameter LSTM) kestabilan pelatihan, khususnya pada hidden size besar. Meskipun latensi inferensi keduanya serupa, SSM terbukti lebih ringan secara komputasi dan mudah dioptimasi, menjadikannya alternatif yang lebih praktis untuk sistem inferensi real-time berskala besar. Temuan ini mengindikasikan bahwa SSM merupakan alternatif yang efisien dan praktis untuk diterapkan dalam sistem inferensi real-time.

.....In the digital era, video has become the primary form of content across various social media platforms such as TikTok. With millions of videos uploaded daily, there is a need for an automated classification system capable of quickly and efficiently processing and categorizing content to support content moderation. This research aims to design and evaluate a real-time inference system based on deep learning for TikTok video classification, focusing on system efficiency, model accuracy, and training stability. The methodology involves modifying the Long-term Recurrent Convolutional Network (LRCN) architecture by replacing the LSTM with a State Space Model (SSM) as a drop-in replacement for modeling temporal sequences. The system architecture consists of a CNN for spatial feature extraction, a 3-layer neural network for intermodule connections, and the use of ZeroMQ and Docker for data sequencing and inference process isolation. Evaluation is conducted under both batch and real-time scenarios, including measurements of accuracy, F1-score, parameter count, inference latency, and training stability. The results show that SSM and LSTM achieve comparable classification accuracy (about 76%), but SSM have fewer parameter (50% of LSTM) and training stability, especially at larger hidden sizes. Although their inference latencies are similar, SSM is computationally lighter and easier to optimize, making it a more practical alternative for large-scale real-time inference systems. These findings indicate that SSM is an efficient and practical alternative for implementation in real-time inference systems.