

# AngularGrad: A New Optimization Technique for Angular Convergence of Neural Networks

## Các tóm tắt chính

Nguyễn Hồng Sơn<sup>\*†</sup>

<sup>\*</sup>Trường Đại học Công nghệ Thông tin

<sup>†</sup>Email: sonnh.17@grad.uit.edu.vn

*Tóm tắt nội dung*—Mạng nơ ron tích chập (CNN) được huấn luyện sử dụng các trình tối ưu hóa dựa trên độ dốc giảm dần ngẫu nhiên (SGD). Gần đây, trình tối ưu hóa ước lượng thời điểm thích ứng (Adam) đã trở nên rất phổ biến do động lực thích ứng của nó, giải quyết vấn đề độ dốc sắp chết của SGD. Tuy nhiên, hiện có trình tối ưu hóa vẫn không thể khai thác độ cong tối ưu hóa thông tin một cách hiệu quả. Bài báo này đề xuất một AngularGrad mới trình tối ưu hóa xem xét hành vi của hướng/góc của gradient liên tiếp. Đây là nỗ lực đầu tiên trong văn học để khai thác thông tin góc gradient ngoài kích cỡ. AngularGrad được đề xuất tạo ra điểm số để kiểm soát kích thước bước dựa trên thông tin góc gradient của các lần lặp trước đó. Vì vậy, các bước tối ưu hóa trở nên mượt mà hơn vì kích thước bước chính xác hơn của quá khứ ngay lập tức độ dốc được ghi lại thông qua thông tin góc. Hai các biến thể của AngularGrad được phát triển dựa trên việc sử dụng Các hàm tang hoặc Cosine để tính toán góc gradient thông tin. Về mặt lý thuyết, AngularGrad cũng thể hiện sự hồi tiếc tương tự bị ràng buộc như Adam cho mục đích hội tụ. Tuy nhiên, rộng rãi các thử nghiệm được thực hiện trên các tập dữ liệu chuẩn so với các phương pháp tiên tiến cho thấy hiệu suất vượt trội của AngularGrad.

*Từ khóa*—

## I. ĐẶT VẤN ĐỀ