

ABSTRAK

Indonesia sebagai negara agraris menghadapi ancaman gagal panen akibat penyakit tanaman, dan keterbatasan petani mengenali gejala pada permukaan daun mendorong pemanfaatan *deep learning* berbasis *Convolutional Neural Network* (CNN). Namun arsitektur CNN *state-of-the-art* umumnya berparameter masif sehingga sulit diterapkan secara *standalone* pada perangkat seluler berspesifikasi terbatas. Belum tersedia pemetaan komprehensif mengenai arsitektur mana yang paling tahan terhadap kompresi ekstrem, bagaimana strategi pra-pelatihan berinteraksi dengan ketahanan tersebut, serta metrik yang menjamin keadilan perbandingan antar-topologi.

Penelitian ini memetakan ketahanan 11 arsitektur CNN yang mewakili variasi topologi—*residual*, *dense*, *inception*, dan *depthwise separable convolution*—di bawah *Iterative Magnitude Pruning* (*sparsity* nominal 90%, penjadwalan *PolynomialDecay*) dan *Post-Training Dynamic Range Quantization*, lintas tiga skenario pra-pelatihan: bobot acak (Skenario 1), *transfer learning* penuh (Skenario 2), dan *ekstraktor* fitur beku (Skenario 3). Hasil utama dilaporkan pada protokol global yang memangkas seluruh lapisan memiliki *kernel*; protokol selektif yang melindungi lapisan *depthwise* dan *separable* ditampilkan terbatas sebagai demonstrasi kerancuan perbandingan. Skenario 1 dan 2 direplikasi tiga *seed* independen. PlantVillage dipakai untuk evaluasi terkontrol, PlantDoc untuk generalisasi lintas-domain. Diformalkan tiga instrumen metodologis: indeks *Pruning Exposure*, indeks *Critical Effective Sparsity* (S^*_{95}) beserta dekomposisinya menjadi komponen intrinsik dan pemulihan, serta protokol *Exposure-Matched Benchmarking*. Ketidaksahihan perbandingan pada *sparsity* nominal yang sama dinyatakan formal sebagai Lemma Kerancuan *Sparsity* Nominal beserta buktinya. Perbedaan performa divalidasi dengan uji non-parametrik.

Hasil mengungkap dikotomi ketahanan yang tegas: 21 dari 22 konfigurasi mempertahankan akurasi sekurang-kurangnya 95% nilai *baseline* meskipun 90% bobot dipangkas, dengan simpangan baku di bawah 1,0 poin pada 19 konfigurasi. NASNetMobile satu-satunya pengecualian, dan replikasi mengungkap arsitektur tersebut bukan konsisten rapuh melainkan tidak dapat diprediksi (simpangan baku 34,42 poin). Uji Wilcoxon menunjukkan strategi pra-pelatihan berpengaruh nyata ($p = 0,0029$) meskipun kecil, sedangkan Kruskal-Wallis tidak menemukan perbedaan nyata antar-kelompok topologi—kesetaraan yang justru menegaskan sifat dikotomis ketahanan. Dekomposisi S^*_{95} mengungkap MobileNetV2 berketahanan intrinsik terendah (30,65%) namun pulih konsisten hingga 97,83% berkat kapasitas pemulihan melalui *fine-tuning*, sehingga kebebasan jaringan menata ulang dirinya selama pemulihan menjadi penentu ketahanan. Kuantisasi INT8 memampatkan ukuran model hingga 4 kali tanpa penurunan akurasi berarti. Sebagai sintesis dirumuskan Skor Kesiapan Kompresi (CRS) beserta kriteria kesiapan dua syarat yang mempertimbangkan besaran sekaligus kestabilan ketahanan. Temuan ini memberikan panduan pemilihan arsitektur dan strategi pra-pelatihan untuk pengerahan efisien di perangkat seluler pada domain terkontrol, sekaligus menyumbangkan kerangka metrik yang lebih adil bagi perbandingan kompresi lintas-topologi.

Kata kunci: *deep learning*, penyakit tanaman, pemangkasan (*pruning*), kuantisasi, ketahanan arsitektur, *pruning exposure*, *critical effective sparsity*, generalisasi lintas-domain.

ABSTRACT

Indonesia, an agrarian country, faces crop failure caused by plant diseases, while farmers' limited ability to recognise leaf symptoms has encouraged the adoption of deep learning based on Convolutional Neural Networks (CNN). State-of-the-art CNN architectures, however, are massively parameterised and difficult to deploy standalone on resource-constrained mobile devices. No comprehensive mapping is yet available regarding which architectures are most resilient to extreme compression, how pre-training strategies interact with that resilience, or which metrics guarantee fair cross-topology comparison.

*This study maps the resilience of eleven CNN architectures spanning residual, dense, inception, and depthwise separable topologies under Iterative Magnitude Pruning (90% nominal sparsity, PolynomialDecay scheduling) and Post-Training Dynamic Range Quantization, across three pre-training scenarios: random weights, full transfer learning, and frozen feature extraction. Results are reported under the global protocol, which prunes all kernel-bearing layers; the selective protocol, protecting depthwise and separable layers, is presented in limited scope to demonstrate comparison confounding. The first two scenarios were replicated using three independent seeds, with PlantVillage for controlled evaluation and PlantDoc for cross-domain generalisation. Three methodological instruments are formalised: the Pruning Exposure index, the Critical Effective Sparsity index (S^*95) decomposed into intrinsic and recovery components, and the Exposure-Matched Benchmarking protocol. The invalidity of comparison at identical nominal sparsity is stated formally as the Nominal Sparsity Confounding Lemma with its proof. Differences were validated non-parametrically.*

*A sharp resilience dichotomy emerged: twenty-one of twenty-two configurations retained at least 95% of baseline accuracy despite 90% of weights being pruned, with standard deviations below 1.0 point in nineteen configurations. NASNetMobile was the sole exception, and replication revealed it to be not consistently fragile but unpredictable (standard deviation 34.42 points). Wilcoxon showed a significant yet small effect of pre-training strategy ($p = 0.0029$), whereas Kruskal-Wallis found no difference between topological groups—an equivalence reinforcing that dichotomy. Decomposition of S^*95 revealed that MobileNetV2 has the lowest intrinsic resilience (30.65%) yet recovers consistently to 97.83% through fine-tuning, establishing a network's freedom to reorganise during recovery as the decisive determinant of resilience. INT8 quantization compressed model size up to fourfold without appreciable accuracy loss. A Compression Readiness Score (CRS) is then formulated with a two-condition readiness criterion accounting for both the magnitude and stability of resilience. These findings guide architecture and pre-training selection for efficient mobile deployment within a controlled domain, and contribute a fairer metric framework for cross-topology compression comparison.*

Keywords: *deep learning, plant disease, pruning, quantization, architectural resilience, pruning exposure, critical effective sparsity, cross-domain generalisation*