

# TS2 Test Senaryosu - Özet Rapor

## Test Konfigürasyonu

### Model ve Veri Seti

- **Model:** ResNet50
- **Veri Seti:** ImageNet
- **Validasyon Örnekleri:** 49,997
- **Kalibrasyon Örnekleri:** 25,600

### Pruning Parametreleri

- **Pruning Oranı:** 20%
- **Global Pruning:** Hayır (Layer-wise)
- **İteratif Adımlar:** 1
- **Fine-Tuning Epoch:** 5
- **Kalibrasyon Batch:** 100

### Test Edilen Yöntemler

1. **Neuron Coverage Pruning** - Nöron aktivasyon kapsama tabanlı budama
2. **Wanda Pruning** - Ağırlık  $\times$  Aktivasyon önem tabanlı budama

## Karşılaştırmalı Sonuçlar

Yöntem	Doğruluk (%)	Doğruluk Kaybı	Boyut (MB)	Boyut Azalması	Çıkarım Süresi (ms)	FLOPs (G)	FLOPs Azalması
<b>Orijinal (Fine-Tuned)</b>	78.59	-	97.70	-	0.029	4.13	-
<b>Coverage Pruning</b>	48.44	-30.15	63.64	-34.9%	1.19	2.66	-35.7%
<b>Wanda Pruning</b>	56.45	-22.14	63.64	-34.9%	1.19	2.66	-35.7%

## Önemli Gözlemler

### Performans Karşılaştırması

- **Wanda Pruning** yöntemi, Coverage Pruning'e göre **8.01% daha yüksek doğruluk** sağlamıştır
- Her iki yöntem de aynı model boyutu ve FLOPs azalmasına ulaşmıştır

- Wanda yöntemi, aktivasyon ve ağırlık önemini birleştirerek daha iyi performans göstermiştir

### **Model Optimizasyonu**

- **%34.9 model boyutu azalması** (97.70 MB  $\rightarrow$  63.64 MB)
- **%35.7 FLOPs azalması** (4.13G  $\rightarrow$  2.66G)
- Her iki yöntemde de çıkarım süresi artmıştır ( $\sim 0.03\text{ms} \rightarrow \sim 1.19\text{ms}$ )

### **Doğruluk-Verimlilik Dengesi**

- Coverage Pruning: Daha agresif budama, %38.4 orijinal performans korundu
- Wanda Pruning: Daha dengeli budama, %71.8 orijinal performans korundu