

## Abstrak

Produksi pisang di Indonesia pada tahun 2022 mencapai 9,6 juta ton buah. Metode konvensional yang digunakan untuk menentukan tingkat kematangan pisang masih mengandalkan indera penglihatan manusia dengan memperhatikan perubahan warna kulit pisang. Namun, penentuan tingkat kematangan pisang dengan metode ini memiliki beberapa kekurangan, seperti waktu yang lama dan bersifat subjektif sehingga dapat menghasilkan penilaian yang berbeda-beda pada setiap individu. Oleh karena itu, teknologi *computer vision* dapat menjadi solusi yang efektif dalam mengklasifikasikan kematangan buah pisang secara otomatis. Penelitian ini menggunakan metodologi Vision Transformer (ViT) untuk mengklasifikasikan tingkat kematangan pada pisang, dengan tingkatan yang terbagi ke dalam empat kelas, yaitu mentah, semi-matang, matang, dan sangat-matang. Penelitian dilakukan dengan menggunakan lima model ViT yang sudah dilatih sebelumnya atau *pre-trained*, yaitu ViT-B-P16, ViT-B-P32, ViT-L-P16, ViT-L-P32, dan ViT-H-P14 pada ImageNet-21k dan ImageNet-1k. Kemudian, model ViT tersebut dievaluasi dan dibandingkan dengan model CNN. Evaluasi dilakukan menggunakan dataset uji dengan 5.068 citra pisang hasil penggabungan dari dataset yang dipublikasikan secara *online*. Hasil evaluasi menunjukkan model ViT-L-P16-in21k memiliki akurasi tertinggi sebesar 91,61%. Model ViT menunjukkan kemampuan generalisasi yang lebih baik, sementara CNN memiliki ukuran model dan waktu pelatihan yang lebih efisien.

**Kata Kunci:** klasifikasi kematangan pisang, *computer vision*, vision transformer, *pre-trained model*