An Integrated Pedagogical Approach to Teaching Convolution Filtering and Principal Component Analysis

Abbas Attarwala*
California State University, Chico
Chico, CA, USA
Aattarwala@csuchico.edu

Ed Lindoo
Regis University
Denver, CO, USA
elindoo@regis.edu

Abstract

In many undergraduate programs, Principal Component Analysis (PCA) is introduced in applied linear algebra or data science courses as a standalone method for dimensionality reduction. Separately, students may encounter the discrete convolution in courses on systems or signal processing, again without explicit connection to PCA. Although some advanced courses in digital image processing may revisit both topics, our goal is different: to deliberately integrate these concepts within an applied linear algebra course to explicitly demonstrate their interaction. This experience paper describes how combining discrete convolution-based filtering (e.g., low-pass and high-pass) with PCA enables students to investigate how preprocessing decisions shape eigenvalue distributions, variance concentration, and finally the interpretation of data. By treating these methods as interdependent components of a data pipeline rather than isolated tools, we believe that students gain a deeper understanding of the subjective assumptions embedded in the analysis. This integrated approach encourages students to engage with linear algebra as a context-sensitive discipline where choices matter.

CCS Concepts

• **Computing methodologies**; • **Mathematics of computing**; • **Applied computing**; • **Social and professional topics**;

Keywords

Convolution Based Filtering, Principal Component Analysis, Pedagogical Teaching, Signal, Noise

ACM Reference Format:
Abbas Attarwala and Ed Lindoo. 2025. An Integrated Pedagogical Approach to Teaching Convolution Filtering and Principal Component Analysis. In *26th Annual ACM Conference on Cybersecurity and Information Technology Education (ACM SIGCITE 2025), November 06–08, 2025, Sacramento, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3769694.3771149

1 Introduction

In typical undergraduate programs, Principal Component Analysis (PCA) and Discrete Convolution (DC) are introduced in separate courses and disciplinary contexts. For example, electrical and computer engineering (ECE) students can learn about DC in systems and signal analysis courses [7, 12], while ECE and computer science (CS) students can study PCA in an applied linear algebra course [8–10] and data science students can examine it in an introductory machine learning course [1]. Both PCA and DC-based filtering appear in some upper-level courses, such as Colorado State University's ECE 513: Digital Image Processing [3]. Although it is certainly possible for instructors to make conceptual connections between these topics during instruction, the publicly available syllabus presents them as part of separate modules. This kind of modular structure is common across many course syllabi and may limit opportunities for students to explore how these techniques influence each other.

DC often goes untaught in CS programs, appearing primarily in signals and systems courses that many CS students never take. As a result, PCA is frequently treated as an isolated, and most likely assumption-free technique for dimensionality reduction. This obscures a critical fact: preprocessing with DC-based filtering determines what PCA treats as signal versus noise, fundamentally shaping its output.

To address this gap, we present a pedagogical approach that integrates PCA and DC in an applied linear algebra course for CS students. By introducing these concepts together rather than in isolation, we help students develop a deeper understanding and critical reasoning about data transformations. Rather than treating this integration as an added topic, we frame it as a structured learning sequence within applied linear algebra: introducing DC as a mathematical model of system response, applying it in the context of digital filtering (e.g., low-pass and high-pass), and then analyzing how these filters affect PCA's interpretation of data. Students learn how filtering redistributes variance across principal components, fundamentally altering which eigenvectors capture the most variance. This technical understanding leads directly to an important realization: What constitutes signal versus noise depends on the task-specific priorities. Consider detecting fruit from aerial images. A high-pass filter emphasizes object boundaries, potentially aiding detection, while a low-pass filter suppresses high-frequency details like edges and texture, creating smoother images that compress more efficiently, but may obscure features critical for detection. By applying PCA after each filtering choice, students observe that high-pass filtered images require many more principal components to capture the same variance as low-pass filtered images. This quantitative comparison reveals how the preprocessing choice directly impacts the dimensionality requirements and which approach better supports their specific goal.

This hands-on exploration reveals that filtering decisions are not neutral, but encode assumptions about what matters in the data. Students practice hypothesis formation (What will happen if I apply this filter?), quantitative validation (examining eigenvalue distributions), and reflective evaluation (assessing whether a

ACM SIGCITE 2025, Sacramento, CA, USA

ACM ISBN 979-8-4007-2240-0/25/11
https://doi.org/10.1145/3769694.3771149

transformation serves the task). These skills transfer broadly across domains: from anomaly detection in computer security to feature extraction in data science to log preprocessing in information technology. Most importantly, students learn to embrace and navigate the inherent ambiguity in real-world data analysis. Rather than seeking a correct preprocessing step, they learn to make informed, context-aware decisions about data transformations, developing the critical thinking habits essential for any data-driven field.

The remainder of this paper is organized as follows. Section 2 reviews related work on educational approaches to PCA and filtering. Section 3 presents our integrated pedagogical framework and lesson structure. Section 4 and Section 5 present future work and conclusions.

2 Literature Review

Although our classroom use of PCA focuses on visual reasoning and student prediction, similar principles of dimensionality reduction and feature extraction appear throughout real-world deep learning systems. While some researchers [6] have combined PCA with deep learning architectures (using PCA to weight CNN feature vectors for improved classification), the pedagogical value of teaching how convolution-based filtering affects PCA's interpretation of data remains largely unexplored.

[15] investigated feature selection mechanisms for classroom sound classification, proposing a PCA-based feature ranking method and comparing it to the Relief-F algorithm. Their work focused on identifying optimal subsets of features from a set of 143 temporal, spectral, and perceptual descriptors to improve classification accuracy using machine learning models such as LDA, QSVM, kNN, Boosted Trees, and Random Forest. Although their study demonstrates the utility of PCA in efficiently ranking features to support classification accuracy, it is primarily algorithmic in nature. In contrast, our work focuses on the pedagogical value of PCA and filtering as tools to build the geometric intuition of students. Rather than evaluating PCA utility based on accuracy metrics in classification models, we use simple, visual data sets to help students reason about how high-frequency disruptions affect variance alignment and how low-pass filtering can help restore structure that PCA can then efficiently capture. Our approach is not aimed at model optimization but at creating conceptual understanding for our students about how PCA operates under different data conditions.

[14] present an effective pedagogical approach to teaching PCA through hands-on activities using real-world datasets and open-source tools. By integrating the R programming language and its RCommander graphical interface plugin, they make PCA accessible to students with limited programming experience. Their curriculum emphasizes the interpretation of score and loading plots, exploration of variable importance, and the effect of data preprocessing using multiple chemometric case studies. Students learn to recognize patterns, outliers, and correlations in a range of real-world datasets, including elemental properties, food composition, and biomedical data. In contrast, our work emphasizes a geometric and variance-focused interpretation of PCA through carefully designed synthetic datasets. Instead of relying on domain-specific data, we use controlled examples such as a $x = y$ line with localized high-frequency disruptions to help students build visual and conceptual intuition. This allows learners to see, in real time, how off-axis variance affects the alignment of principal components and how tools like low-pass filters can restore that alignment. While [14] guide students in applying PCA to complex datasets, our approach scaffolds student understanding from visual pattern recognition to more abstract understanding of how PCA partitions variance and responds to preprocessing.

[11] offer a comprehensive systematic review of 82 peer-reviewed studies that take advantage of audio features in educational research, covering work from 2014 to 2024. The review identifies three core categories of audio-derived data: low-level acoustic characteristics (e.g. MFCCs, pitch), diarization-based indicators (e.g. speaker turn-taking, talk ratios), and linguistic characteristics derived from transcripts, and examines how these are extracted, combined, and applied across a range of educational contexts. The authors highlight both technical progress and systemic limitations, including limited user feedback, poor model interpretability, and the absence of shareable anonymized data sets. Their analysis presents two divergent research workflows: one focused on technical modeling and the other on pedagogical relevance, arguing that true educational impact depends on bridging this gap. To this end, they advocate for the release of standardized feature-level datasets, the participatory design of feedback systems with educators, and the integration of generative AI to translate analytics into actionable, context-aware instructional guidance.

Although [11] emphasize the importance of making audio analytics pedagogically relevant, their review focuses primarily on technical pipelines and large-scale system design. In contrast, our work takes a bottom-up, student-centered approach, using controlled visual examples to build an intuitive understanding of how PCA interacts with filtered and unfiltered data. We focus not on a general analytics infrastructure but on cultivating the geometric reasoning of individual learners about variance, alignment, and reduction in dimensionality through hands-on experiments and visualization.

[5] provides a comprehensive AI-assisted survey of pedagogical practices in DSP education, covering a range of teaching methods, from traditional lectures to interactive and technology-enhanced strategies. The review emphasizes the effectiveness of flipped classrooms, hands-on laboratories, simulation tools (e.g. MATLAB, FPGA), and project-based learning in improving student comprehension and practical skill development. It also advocates for AI-supported teaching aids and remote lab access as a means of democratizing DSP education. Through systematic synthesis, the paper highlights the importance of aligning DSP instruction with constructivist pedagogies and evolving industry needs. While [5] focuses on general pedagogical strategies for DSP, our work focuses on the conceptual and visual intuition behind PCA and discrete convolution filtering, specifically emphasizing how high and low-pass filtering affects principal component alignment and structure recovery. Rather than a literature review or survey, we contribute a guided concept-first approach with visual demonstrations that help students internalize the role of frequency-domain filtering in signal decomposition and compression.

[4] observed that the application of low-pass filters to the image data significantly improved the PCA results by increasing the relative variance captured by the initial principal components. However,

in the MATLAB forum, there is no other response that explains this observation. However, in our paper, we present a pedagogical approach that not only explains this but also presents a teaching approach to introduce such material to CS students taking an applied linear algebra course.

[16] explored the practical applications of linear algebra and statistics in image processing, focusing on two fundamental techniques: Gaussian filtering and PCA. [16] established that image processing, beyond just editing, is crucial to pre-processing data in machine learning-based image recognition. The study specifically details how Gaussian filters leverage linear transformations to effectively remove noise from digital images, demonstrating their superiority over simpler averaging filters in preserving image details due to their weighted averaging based on a Gaussian distribution. Furthermore, [16] elucidates the role of PCA in image processing, explaining how it utilizes the statistical information of observed data to extract significant features while simultaneously reducing the data size. [16] delves into the mathematical underpinnings of PCA, including the computation of mean vectors and covariance matrices, and the identification of principal components through eigenvalues and eigenvectors. This reduction in dimensionality is critical to efficiently processing massive datasets inherent in digital images without a significant loss of detail. [16] also highlights an interesting application of PCA in facial recognition, known as eigenfaces, where new images are projected onto a subspace spanned by these principal components for classification.

Although the existing literature demonstrates the technical benefits of combining filtering with PCA, no prior work addresses the fundamental educational challenge of helping students understand how preprocessing decisions embed assumptions and propagate through analytical pipelines. Current approaches treat PCA and convolution filtering as independent topics, failing to develop critical thinking about analytical choices and competing objectives. Students learn procedural knowledge without understanding that mathematical tools embody specific assumptions about data structure, leading to passive tool application rather than informed decision-making.

Our integrated pedagogical framework addresses this gap by teaching students to recognize that preprocessing choices fundamentally transform what PCA considers important in data. Unlike traditional approaches that emphasize technical proficiency with isolated tools, our method employs progressive discovery experiences and prediction-based learning to develop critical thinking about the interconnected nature of analytical decisions, preparing students for real-world scenarios where domain expertise and objective specification drive preprocessing choices.

3 Pedagogical Framework: Integrating PCA and Convolution Through Progressive Discovery

This section presents our integrated pedagogical approach, which guides students from an intuitive visual understanding of PCA and DC to a deeper grasp of how early processing decisions shape the results of data analysis. We implement this approach through interactive MATLAB demonstrations, leveraging the platform's visualization capabilities to make abstract concepts tangible. The framework employs a predict-pause-reveal methodology that engages students in hypothesis formation and empirical testing, transforming them from passive recipients of mathematical procedures into critical thinkers about analytical choices.

3.1 Building Intuition: Visual PCA Understanding Without Mathematics

We begin with three intentionally designed 2D toy datasets that help students develop geometric intuition about principal components before encountering any mathematical formalism. Each data set as shown in Figure 1 is constructed so that the data is visually spread along a clear direction such as the x axis,y axis, or the diagonal line $y = x$, which makes it easy for students to reason about where the variance is concentrated. Students are asked to examine the scatter plots and sketch, on paper, a vector they believe points in the direction of the maximum spread of the data. In small groups, they spend 2–3 minutes discussing and recording their predictions about the direction of the principal components. The visual clarity of these data sets makes the task intuitive: for data spread along the x axis, students naturally draw their vector horizontally; for data along the y axis, they draw it vertically; and for data along the diagonal $y = x$ they sketch a diagonal vector. This immediate visual feedback builds confidence and geometric understanding. Because the data sets are visually interpretable, students can engage in meaningful hypothesis generation without having to compute a covariance matrix, even though they have already been introduced to computing eigenvectors and eigenvalues numerically. This low-stakes collaborative activity builds confidence in their visual reasoning while naturally setting up the need for formal tools—specifically, the covariance matrix and eigenvalue decomposition—to test their intuition. The pedagogical approach draws on our previous work that emphasizes the introduction of PCA through geometric and experiential learning pathways [2]. After the students share their predictions, we compute and reveal the principal components using `eig(cov(Data))` in MATLAB. This moment becomes a powerful pivot: Students realize that the vectors they sketched are, in fact, the eigenvectors of the covariance matrix, the principal components, giving concrete meaning to a concept that is often introduced abstractly. This visual first approach aligns with the pedagogical strategies for eigen theory that emphasize the building of embodied understanding before formal mathematical treatment [13].

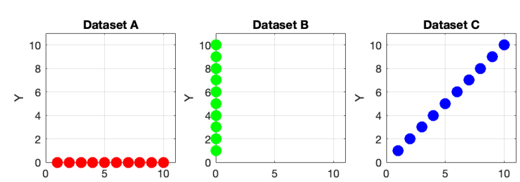

Figure 1: Three fundamental datasets for building PCA intuition. Students predict principal component directions before mathematical analysis: (a) X-axis data, (b) Y-axis data, (c) Diagonal data.

After identifying principal components in clean data as seen earlier in Figure 1, students now examine how robust these components are to noise. They predict how random variations may affect both the direction of the principal component and the concentration of variance. We pose two key questions: (1) Will the direction of the principal component change significantly? and (2) Will the first component still capture the same percentage of the total variance? We then introduce small random variations to the data to test whether a single vector can still capture the complete spread of the dataset. In Figure 2, students observe that with noise present, they now need two principal components to capture all variance: no single vector can account for the entire variance due to random variations. Computing the eigenvalues and eigenvectors using MATLAB's `eig(cov(Data))` function on the noise-perturbed dataset reveals that even small random perturbations reduce the variance captured by the first principal component from 100% to about 98%. This slight decline has geometric significance: the variation introduced by noise is not perfectly aligned with the original dominant direction. As a result, a single vector no longer suffices to capture the entire spread of the data, and a second component becomes necessary to account for the remaining variance.

This observation motivates our next exploration: how filtering can increase the variance captured by the first principal component. While the first component remains dominant even with noise (capturing 98% of variance), filtering can potentially restore its ability to capture nearly all the variance by removing the scattered contributions of noise. In the following sections, we explore how frequency-based filtering through convolution affects the concentration of variance in principal components, examining this relationship through structured signal patterns.

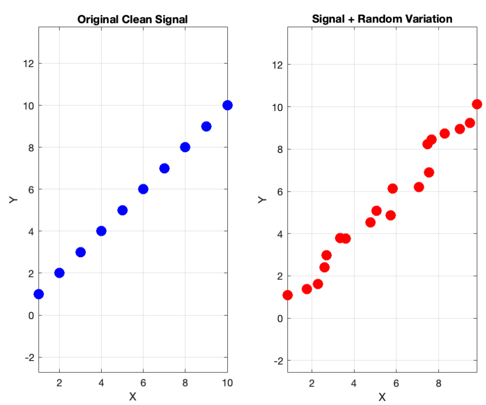

Figure 2: Effect of adding random variations to pure diagonal signal. (a) Clean signal with 100% variance in first component, (b) Signal plus noise with reduced variance concentration.

3.2 Filter Concepts: Localized High-Frequency Disruption

To introduce convolution or filtering as a PCA preprocessing tool, we present students with a structured data set that begins with a clean linear relationship ($x = y$), then selectively disrupt it by adding high-frequency variation to a small portion of the data. The pedagogical goal is for students to see, both visually and statistically, how this localized disruption misaligns part of the dataset from the dominant trend and how filtering can restore structure in a way that is meaningful to PCA.

3.3 Localized Noise and Filtering to Restore Structure

At this point, we pose the following question to students: *"If we add rapid fluctuations to only part of the dataset, how will that affect the direction of the principal component and the amount of variance it captures? What happens if we smooth these fluctuations out again?"* To explore this question, students work with a two-dimensional dataset constructed as follows: the x values increase linearly, while the y values follow $y = x$ plus a smooth low-frequency sinusoidal variation applied to all points. We then inject high-frequency oscillations into a localized region as seen in Figure 3 in the left image, creating sharp fluctuations in y within that window (from x=4 to x=6). This produces a visible distortion: a localized "burst" of rapid oscillations that deviates from the otherwise smooth global trend.

The low-frequency variation results in gentle and consistent changes, whereas the high-frequency component introduces abrupt and localized variation. This setup makes it easy for students to distinguish between smooth global structure and targeted noise.

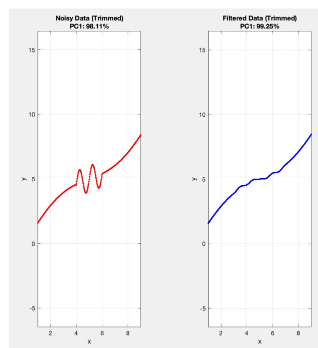

Figure 3: Localized high-frequency variation disrupts the $x = y$ structure, visibly misaligning the data. After low-pass filtering, the dominant linear trend is restored, allowing PCA to capture more variance with a single component.

When presented with this visualization, students often describe the high-frequency burst as an anomaly. They recognize that this localized disruption introduces uncorrelated behavior that pulls part of the data away from the primary direction. As a result, the first principal component still captures the dominant direction of variation, but less efficiently; some variance is pushed into secondary components due to localized disruption. This observation leads to a crucial pedagogical moment: we demonstrate that a low-pass filter specifically targets and removes high-frequency oscillations while preserving the underlying smooth structure. Students see why we choose a low-pass filter rather than a high-pass filter: since the anomaly was created by adding high-frequency components, a low-pass filter acts as the appropriate antidote, allowing only the smooth, low-frequency variations to pass through while eliminating the rapid oscillations. A single principal component once again explains a larger share of the variance (99.25% after filtering, as shown on the right in Figure 3, compared to 98.11% before filtering on the left). This concrete demonstration reinforces the mathematical relationship: high-frequency noise scatters variance across multiple components, while low-pass filtering consolidates it back into the primary direction. This connection between visual structure, filter selection, PCA behavior, and variance concentration deepens students' understanding of how thoughtful data preprocessing directly impacts dimensionality reduction and interpretability.

3.4 Connecting to Real-World Applications

After demonstrating with simple 2D examples that low-pass filtering concentrates variance in the first principal component, we now apply this principle to realistic scenarios where preprocessing choices have practical consequences. We introduce two contrasting applications that show how the same pre-processing technique can be optimal for one objective while counterproductive for another. This demonstrates that analytical choices are not universally correct or incorrect, but must be justified by the specific context and goals.

3.4.1 Image Compression: When Efficiency Matters Most. We present students with a realistic image compression scenario using a grayscale image as seen in Figure 4 that contains both meaningful structural information and high-frequency details.

We introduce the Gaussian filter as a specific type of low-pass filter that smooths images by replacing each pixel's value with a weighted average of its neighboring pixels, where the weights decrease with distance from the center pixel. This is the same principle students saw earlier when we used a low-pass filter to remove the high-frequency anomaly from our dataset: The filter preserves smooth, gradual variations while eliminating rapid oscillations. We ask students to predict how Gaussian filtering will affect PCA compression efficiency based on their understanding from the anomaly removal example and diagonal signal demonstration. Key guiding questions include: *If we apply the same low-pass filtering principle you observed in Figure 3, what should happen to the number of principal components needed to capture 90% of the image variance?* and *What assumptions are we making about which image features are important for compression purposes?* Most students correctly predict improved compression efficiency, reasoning that removing high-frequency variations concentrates variance in fewer components. They also identify the key assumption: natural images have strong spatial correlation (neighboring pixels share similar values), making smooth, low-frequency content more important than fine textures and edges for compression purposes. Figure 4 reveals the quantitative magnitude of this effect. Students observe that the top 30 principal components capture 83.4% of variance in the original image, while the same number of components capture 90.6% of variance after Gaussian filtering.

This demonstration teaches students that preprocessing is a design choice that fundamentally affects system performance. By achieving better compression ratios through filtering before PCA, they see the practical interplay between signal processing and dimensionality reduction. In real-world scenarios with limited storage or bandwidth, this knowledge enables them to represent images with fewer components while maintaining acceptable quality.

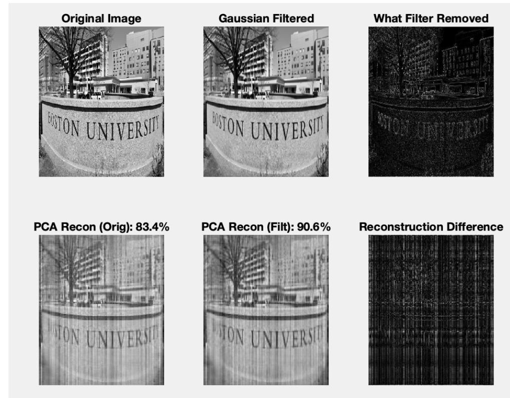

Figure 4: Image compression demonstrating how Gaussian filtering improves PCA efficiency. Filtering increases variance captured by 30 components from 83.4% to 90.6%.

After observing these results, we challenge students to articulate what assumptions the Gaussian filter encoded about the image content. We facilitate this through specific prompts: *What visual features become blurred or softened when you compare the filtered and unfiltered images?* Students typically identify smoothing of texture details, softening of sharp boundaries, and reduction of fine-grained patterns. We then pose a critical question: *If you were compressing a medical image versus a landscape photo, would you make the same filtering choice?* This helps students recognize that filtering involves trade-offs: improved compression efficiency comes at the cost of high-frequency information that might be crucial for certain applications. Medical diagnosis might require preserving fine details that artistic photography could safely smooth out. Through this discussion, we hope students understand that filtering is not merely a technical pre-processing step but an implicit decision about what information matters for the specific task at hand.

3.4.2 Fruit Detection: When Detail Preservation Becomes Critical. Having seen how low-pass filtering optimizes for compression efficiency, students encounter a contrasting scenario where the same preprocessing strategy proves counterproductive. We present a fruit detection challenge using a synthetic canopy image where small fruits must be distinguished from background foliage. Before demonstrating any filtering approaches, we ask students to identify what image characteristics would be most important for automated fruit detection. Through discussion, they typically recognize that fruit boundaries, shape differences, and local intensity variations are crucial features that distinguish fruits from leaves and background elements. We then pose the central question that reveals the conflict between competing objectives: *Given what you know about how Gaussian filtering affects PCA, and what you have identified as important for fruit detection, would you expect the same preprocessing that optimized compression to be optimal for detection?* Students engage in structured hypothesis formation, predicting how different filtering strategies will affect both PCA efficiency and detection accuracy. Most students correctly anticipate that high frequency data is more valuable here to detect the changes in boundaries of fruit and not the low frequency data. Any edge detection filters will better preserve the boundary information crucial for fruit identification, even if this comes at the cost of reduced PCA compression efficiency. The empirical results as seen in Figure 5 validate student predictions while quantifying the tradeoffs involved. Edge detection and high-pass filtering preserve the fine details essential for distinguishing fruits from background, but require more principal components to capture equivalent variance. To understand this tradeoff quantitatively, we examine how different filters affect the distribution of variance across principal components. In our analysis, we compute 30 principal components (which together capture 100% of the variance) and order them by decreasing variance explained—component 30 captures the most variance, component 29 the next most, down to component 1 which captures the least. The bottom panel of Figure 5 reveals a striking contrast: when using edge detection or high-pass filtering, the first 25 components (those explaining the least variance individually) cumulatively capture about 70% of the total variance. This indicates that high-frequency information is distributed across many dimensions, requiring numerous components to represent it adequately. Conversely, after Gaussian (low-pass) filtering, these same 25 components capture only 5% of the variance—meaning 95% is concentrated in just components 26-30. This demonstrates a fundamental principle: low-pass filtering concentrates variance in a few dominant components (ideal for compression), while high-pass filtering spreads variance across many components (necessary for preserving edge details critical to detection). Students discover that preprocessing doesn't just change how much total variance is captured—it fundamentally reshapes where that variance is distributed across the principal component hierarchy. This redistribution determines whether the analysis pipeline will excel at compression (variance concentrated in few components) or feature detection (variance preserved across many components), but rarely both simultaneously.

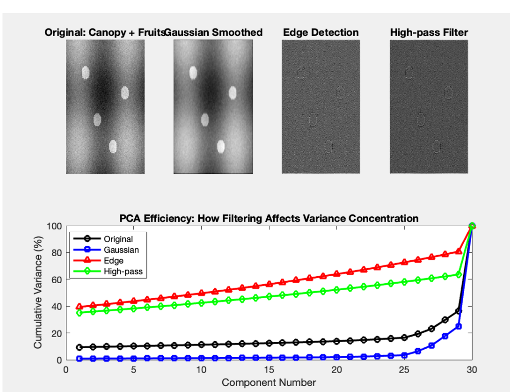

Figure 5: Fruit detection scenario showing tradeoffs between compression efficiency and feature preservation. Different filtering strategies reveal competing preprocessing objectives for detection versus compression tasks.

Students confront the central pedagogical insight through reflection on their empirical observations. We ask: *Can you design a single preprocessing approach that simultaneously optimizes for both compression efficiency and detection accuracy?* Through attempted solutions and subsequent discussion, students discover that such universal optimization is generally impossible, leading to the realization that all analytical choices involve tradeoffs.

4 Future Work

As future work, we aim to more rigorously assess the effectiveness of this integrated pedagogical approach. We plan to design and implement a pre- and post-instruction survey that includes both quantitative Likert scale items (e.g., self-assessed confidence and conceptual clarity) and qualitative open-ended questions (e.g., how students describe the role of filtering in PCA or whether visualizations helped clarify the material). These instruments will be used to gauge how students' understanding evolves during the course and to capture their perceptions of the conceptual connections between PCA and DC filtering. To quantitatively measure the impact of this approach, we also plan to compare learning outcomes across semesters using Welch's t-test, which accommodates unequal variances and sample sizes. Specifically, we will compare aggregated teaching course evaluations results or assessment scores from semesters where PCA was taught alongside DC filtering versus semesters where PCA was taught as a standalone topic. This statistical analysis will allow us to assess whether the integrated approach yields a significant improvement in student understanding, providing evidence to support or refine our pedagogical model.

5 Conclusion

In this paper, we presented an experience-driven exploration of teaching PCA in conjunction with DC-based filtering, specifically using high and low pass filters, to help students build an intuitive understanding of how noise and signal characteristics impact principal component alignment and variance capture. Rather than treating PCA and DC as separate topics, we introduced them as complementary concepts, with the goal of deepening student comprehension of both dimensionality reduction and signal filtering through side-by-side visual demonstrations.

These real-world applications culminate in explicit recognition of how individual decisions propagate throughout analytical workflows. Students trace the logical chain: filter selection determines which image features are preserved, affecting variance distribution across principal components and ultimately influencing compression efficiency and detection effectiveness. This interconnected framework helps students understand that data analysis involves a series of interdependent decisions rather than isolated procedures, making domain expertise essential for effective practice.

Acknowledgments

We gratefully acknowledge the use of OpenAI's ChatGPT for proofreading, grammatical checks, and other text editing tasks.

References

[1] 2020. CS 4786/5786: Machine Learning for Data Science, Spring 2020. https://www.cs.cornell.edu/courses/cs4786/2020sp/. Course webpage—includes description, logistics, topics, assignments, and resources.

[2] Abbas Attarwala. 2024. Rethinking Linear Algebra for Computer Science: Applying Vygotsky's Theory of Learning. *Journal of Computing Sciences in Colleges* 39, 10 (2024), 23–32.

[3] Mahmood R. Azimi. 2025. ECE513: Digital Image Processing. https://www.engr.colostate.edu/ece/wp-content/uploads/2021/10/ECE311_syllabus_2024FA_Pezeshki.pdf Syllabus, Spring 2025.

[4] M. Facchin. 2023. Why does filtering data before PCA improve results? MATLAB Answers. https://www.mathworks.com/matlabcentral/answers/1772560-why-does-filtering-data-before-pca-improve-results Retrieved from https://www.mathworks.com/matlabcentral/answers/1772560-why-does-filtering-data-before-pca-improve-results.

[5] Miguel Bordallo López. 2025. Evolving Pedagogy in Digital Signal Processing Education: AI-Assisted Review and Analysis. *IEEE Access* 13 (2025), 45559–45567.

[6] Amir Mehrabinezhad, Mohammad Teshnehlab, and Arash Sharifi. 2024. A comparative study to examine principal component analysis and kernel principal component analysis-based weighting layer for convolutional neural networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 12, 1 (2024), 2379526.

[7] Department of Electrical and Computer Engineering. 2024. ECE4330: Systems and Signals. https://neuron.eng.wayne.edu/auth/ece4330/ece4330_syllabus.html Syllabus.

[8] Department of Mathematics. 2017. MATH 225: Linear Algebra and Linear Differential Equations. https://web-app.usc.edu/soc/syllabus/20171/30556.pdf Syllabus, Spring 2017.

[9] Department of Mathematics. 2021. MATH 415: Applied Linear Algebra. https://math.illinois.edu/system/files/2021-01/MATH415PL1tentative.pdf Tentative syllabus.

[10] Department of Mathematics. 2024. MAT 167: Applied Linear Algebra. https://www.math.ucdavis.edu/courses/syllabus_detail?cm_id=114 Course syllabus.

[11] Federico Pardo, Óscar Cánovas, and Félix J García Clemente. 2025. Audio Features in Education: A Systematic Review of Computational Applications and Research Gaps. *Applied Sciences* 15, 12 (2025), 6911.

[12] Ali Pezeshki. 2024. ECE311: Linear System Analysis. https://www.engr.colostate.edu/ece/wp-content/uploads/2021/10/ECE311_syllabus_2024FA_Pezeshki.pdf Syllabus, Fall 2024.

[13] Farzad Radmehr. 2024. Learning Eigenvalues and Eigenvectors with Online YouTube Resources: A Journey in the Embodied, Proceptual-symbolic, and Formal Worlds of Mathematics. *PRIMUS* (2024), 1–20.

[14] Laís Feltrin Sidou and Marcel Borges Endler. 2020. Teaching principal component analysis using a free and open source software program and exercises applying PCA to real-world examples. *Journal of Chemical Education* 97, 6 (2020), 1666–1676.

[15] Eleni Tsalera, Andreas Papadakis, and Maria Samarakou. 2021. Novel principal component analysis-based feature selection mechanism for classroom sound classification. *Computational Intelligence* 37, 4 (2021), 1827–1843.

[16] Y. Ukawa. 2025. Image Processing: Gaussian Filtering and Principal Component Analysis. Digital Commons@SHU. https://digitalcommons.sacredheart.edu/acadfest/2025/all/180